

HETEROTRIMERIC G PROTEIN BETA: GAMMA BOUND TO A BIOLOGICALLY
ACTIVE PEPTIDE: STRUCTURAL DEFINITION OF A PREFERRED PROTEIN
INTERACTION SURFACE

APPROVED BY SUPERVISORY COMMITTEE

Stephen R. Sprang, Ph. D.

Elliott M. Ross, Ph. D.

Johann Deisenhofer, Ph. D.

Kevin Gardner, Ph. D

DEDICATION

I would like to thank my family, for always supporting me even when they were sure my course of action was wrong; my friends (especially the Trivia Mafia!) who were always available to dispense advice and compassion in equal measure; my thesis committee for supervising my path without inhibiting my progression upon it; and I would especially like to thank my mentor and his lab for welcoming a student into their world, and making a scientist out of me almost without my realizing it.

“I do not know whether it ought to be so, but certainly silly things do cease to be silly if they are done by sensible people in an impudent way.”

-Jane Austen, *Emma*

HETEROTRIMERIC G PROTEIN BETA: GAMMA BOUND TO A BIOLOGICALLY ACTIVE
PEPTIDE: STRUCTURAL DEFINITION OF A PREFERRED PROTEIN INTERACTION
SURFACE

by

TARA LYNNE DAVIS

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December, 2004

Copyright

by

Tara Lynne Davis, 2004

All Rights Reserved

HETEROTRIMERIC G PROTEIN BETA: GAMMA BOUND TO A BIOLOGICALLY ACTIVE
PEPTIDE: STRUCTURAL DEFINITION OF A PREFERRED PROTEIN INTERACTION
SURFACE

Publication No. _____

Tara Lynne Davis, Ph. D.

The University of Texas Southwestern Medical Center at Dallas, 2004

Supervising Professor: Stephen R. Sprang, Ph. D.

Heterotrimeric G proteins transmit information within and throughout eukaryotic cells. These proteins serve crucial functions in diverse signaling pathways, including visual transduction and adenylyl cyclase regulation. Heterotrimers are composed of α , β , and γ subunits. $G\alpha$ is a guanosine triphosphatase (GTPase); $G\beta$ and $G\gamma$ have no intrinsic enzymatic activity and form an obligate heterodimer that participates in independent signaling interactions when released from $G\alpha$. $G\beta\gamma$ signaling partners include $G\alpha$, phosphatidylinositol 3 kinase, phospholipase C β , adenylyl cyclase, G protein receptor kinase, G protein regulated inwardly rectifying channels,

and N and P/Q type calcium channels. The 2.7Å crystal structure of $G\beta_{1\gamma_2}$ bound to a peptide (SIGKAFKILGYPDYD) obtained from a random peptide library has been solved. This peptide inhibits $G\beta\gamma$ mediated activation of phospholipase C β and phosphatidylinositol-3-kinase γ and binds $G\beta$ with sub-micromolar IC_{50} . $G\beta$ is a seven-stranded β -propeller protein containing WD-40 repeats. SIGK forms a helical structure that binds to the “top” face of the $G\beta$ propeller. The peptide binds the same face of $G\beta$ as $G\alpha$, and mutations of $G\beta$ in this region abrogate peptide binding. In addition, the SIGK peptide binds $G\beta$ using residues sampled by other $G\beta\gamma$ target proteins, both in crystal structures and in mutational analyses. No large conformational changes in $G\beta$ are seen upon SIGK binding, and many of the biological effects seen for the SIGK peptide can be explained by simple competition for the top binding surface on $G\beta$. The SIGK peptide delineates a region on the “top” face of $G\beta$ that is functionally a hot spot for $G\beta\gamma$ target binding. However, this face of $G\beta$ contains a mixture of polar and nonpolar contacts at the binding surface that allows it to interact with proteins of diverse sequence and secondary structure in a unique way compared to other hot spot proteins. The SIGK• $G\beta_{1\gamma_2}$ structure provides insight into the way in which $G\beta$ is able to sustain a range of interactions with multiple binding partners.

TABLE OF CONTENTS

Dedication.....	ii
List of Figures, Tables, and Equations.....	xiii
List of Abbreviations.....	xix
Crystallographic Terms and Abbreviations	xxi
CHAPTER ONE: Introduction	23
Heterotrimeric G Proteins: Background and Biological Relevance.....	23
Heterotrimeric G Protein Cycle	23
G Protein α Subunits and their Relationship to $G\beta\gamma$	25
G Protein β : a WD-40 Repeat Protein.....	29
G Protein γ : Formation of an Obligate Heterodimer with $G\beta\gamma$ and Isoform Specificity.....	35
$G\beta\gamma$ Effector Biology	37
Peptides as $G\beta\gamma$ Targets: Theory and Biological Impact Studies.....	40
Heterotrimeric G Proteins: Structural Studies.....	42
Crystal Structures of Active and Inactive Forms of $G\alpha$	42
The Crystal Structure of Unbound $G\beta\gamma$	45
The Crystal Structure of the $G\alpha\bullet\beta\gamma$ Heterotrimer	48
Crystal Structures of G Proteins in Complex with Effectors and Regulators	50
Description of Dissertation Research	55

BIBLIOGRAPHY	59
CHAPTER TWO: Production of $G\beta_1\gamma_2$ and the $G\beta_1\gamma_2$ •SIGK Complex, and Crystallization of the $G\beta_1\gamma_2$ •SIGK Complex.....	65
Materials and Methods	68
Production and Amplification of $G\beta_1$ and $G\gamma_2$ Baculoviral Constructs	68
Expression and Purification of the $G\beta_1\gamma_2$ Heterodimer	69
Production of SIGK Peptide	71
Formation and Crystallization of $G\beta_1\gamma_2$ •SIGK Complex.....	71
Stabilization and Freezing of Crystals.....	71
Crystal Screening and Data Collection using Synchrotron Sources	72
Results and Discussion	72
Expression and Purification of $G\beta_1\gamma_2$	72
Crystallization of the $G\beta_1\gamma_2$ •SIGK Complex.....	80
BIBLIOGRAPHY	85
CHAPTER THREE: Structure Determination of the $G\beta_1\gamma_2$ •SIGK Complex	87
Materials and Methods	87
Data Collection and Data Reduction	87
Identifying the Correct Space Group of the $G\beta_1\gamma_2$ •SIGK Crystals	88
Introduction to Macromolecular Crystallography: The Phase Problem	88
Solving the Phase Problem using Molecular Replacement	90
Using CNS for Molecular Replacement	92

Using AMoRE for Molecular Replacement	96
Using PHASER for Molecular Replacement	98
Model Refinement.....	100
Map Calculation	100
Model Building	101
Model Validation	101
Results and Discussion	102
Identifying the Correct Space Group for $G\beta_{1\gamma_2}\bullet SIGK$ Crystals with PHASER.	102
BIBLIOGRAPHY	110
CHAPTER FOUR: Structural Analysis of the $G\beta_{1\gamma_2}\bullet SIGK$ Complex	112
Materials and Methods	112
Structure Comparison.....	112
Sequence Alignment.....	113
Calculations of Surface Accessibility, RMS Deviations, and B Factors per Atom	113
Software for Creating Molecular Graphics and other Figure Types	113
Mutational Analysis of $G\beta_1$	114
Small Scale Partial Purification of Biotinylated $G\beta_{1\gamma_2}$ ($b\text{-}\beta\gamma$) and $G\beta_{1\gamma_2}$ Mutants	114
Phage ELISA	115
Determination of $G\alpha_{i1}\bullet G\beta_{1\gamma_2}$ Interactions by Flow Cytometry	116

Results and Discussion	116
Overall Structure of the $G\beta_1\gamma_2$ •SIGK Complex and Mutational Analysis of the SIGK Interface	116
Comparison of the $G\beta_1\gamma_2$ •SIGK Complex to Unbound $G\beta_1\gamma_1$	129
Residues in the $G\beta_1$ Binding Site Contact the SIGK Peptide and Signaling Protein Targets	130
SIGK Effects on $G\beta\gamma$ Signaling Pathways	138
Analysis of the Top Face of $G\beta$ as a “Hot Spot”	143
BIBLIOGRAPHY	150
CHAPTER FIVE: Crystallization of $G\alpha_{i1}$ Bound to Fluorophore Nucleotides: First Attempt	153
The use of Fluorophore Nucleotides to Elucidate G Protein Signaling	153
Materials and Methods	159
Expression and Purification of Untagged $G\alpha_{i1}$	159
Exchange of Fluorophore Nucleotide for Endogenous GDP	160
Crystallography of Untagged $G\alpha_{i1}$ •Fluorophore Complex	160
Molecular Replacement	161
Expression and Purification of Selenomethionine Labeled Untagged $G\alpha_{i1}$	161
Crystallography of Selenomethionine Labeled Untagged $G\alpha_{i1}$	163
MAD Data Collection and Phase Calculation	163
Native Data Collection and Data Reduction	166

Model Building and Refinement	167
Results and Discussion	167
Expression, Purification, and Crystallization of Unlabeled and Selenomethionine Labeled Untagged $G\alpha_{i1}$	167
Identification of EF-Tu as the Crystallization Contaminant	172
Structural Analysis of EF-Tu	174
State of the Bound Nucleotide in each EF-Tu Structure	177
Analysis of the Conformational State of each EF-Tu Structure	183
Troubleshooting Strategies for the Second Attempt at Crystallization	183
BIBLIOGRAPHY	185
CHAPTER SIX: Crystallization of $G\alpha_{i1}$ Bound to Fluorophore Nucleotides: Second Attempt	188
Materials and Methods	188
Expression and Purification of GST- $G\alpha_{i1}$ Constructs	188
Modification of Nucleotide Exchange Protocol	190
Results and Discussion	190
Purification of GST- $G\alpha_{i1}$ Constructs	190
Crystallization of GST- $G\alpha_{i1}$ and GST-mod $G\alpha_{i1}$	195
Crystallization of a $G\alpha_{i1}$ •Fluorophore Complex	196
Effect of Ammonium Sulfite Crystallization Condition on Fluorophore Nucleotides	197

Attempted Optimization of Potassium Phosphate Crystallization Condition of the	
$G_{\alpha_{i1}}$ •Fluorophore Complex	200
Crystal Behavior of the $G_{\alpha_{i1}}$ •BODIPY-GTP γ S Complex	201
BIBLIOGRAPHY	203
CHAPTER SEVEN: Conclusions and Future Directions	204
$G_{\beta\gamma}$ Interactions with Peptide/Protein Targets.....	204
VITAE	207

LIST OF FIGURES, TABLES, AND EQUATIONS

Figure 1-1. The heterotrimeric G protein cycle.....	25
Table1-1. Effectors and regulators of G protein α subunits.....	28
Figure 1-2. The canonical structure of the WD-40 repeat	30
Figure 1-3. Structure of a β -propeller and a WD-40 repeat.....	31
Table 1-2. Effectors and regulators of G protein β subunits.....	32
Figure 1-4. Sequence alignment of representative mammalian $G\beta$ isoforms.....	34
Table 1-3. List of $G\beta$ binding residues derived from mutational and structural studies	39
Figure 1-5. Comparison of the basal and active forms of $G\alpha$	44
Figure 1-6. Nucleotide binding pocket interactions in $G\alpha_{i1}$	45
Figure 1-7. Structure of the unbound form of $G\beta_1\gamma_1$	47
Figure 1-8. Structure of the $G\alpha_{i1}\cdot G\beta_1\cdot G\gamma_2$ heterotrimer.....	49
Figure 1-9. Differences in the nucleotide binding pocket between $G\alpha$ in the heterotrimer and GTP-activated $G\alpha$	50
Figure 1-10. $G\alpha_{i1}$ bound to RGS4 shows the interaction of switch regions with signal transduction regulators	51
Figure 1-11. Structure of the $GRK2\cdot G\beta_1\gamma_2$ complex.....	53

Figure 1-12. Structure of the phosducin•G $\beta_1\gamma_1$ complex	54
Figure 2-1. Peptides from phage display.	67
Figure 2-2. Sequences of proteins used in the current project.....	69
Figure 2-3. Flow chart of G $\beta_1\gamma_2$ (C68S) and G $\beta_1\gamma_2$ (wild-type) purifications.....	74
Figure 2-4. Fractions from Ni-NTA, G $\beta_1\gamma_2$ (C68S).....	74
Figure 2-5. Purification from Hitrap Q anion exchange chromatography, G $\beta_1\gamma_2$ (C68S)	75
Figure 2-6. Chromatograph and fractions from Sephadex 200:75, G $\beta_1\gamma_2$ (C68S).....	76
Figure 2-7. Chromatograph and fractions from Ni-NTA, wild type G $\beta_1\gamma_2$	77
Figure 2-8. Chromatograph and fractions from HiTrapQ, G $\beta_1\gamma_2$	78
Figure 2-9. Chromatograph and fractions from Sephadex 200:75, G $\beta_1\gamma_2$	79
Figure 2-10. Initial crystals of the G $\beta_1\gamma_2$ (C68S)•SIGK complex.....	83
Figure 2-11. Initial crystals of the G $\beta_1\gamma_2$ •SIGK complex solubilized in β -OG	83
Figure 2-12. Initial crystals of the G $\beta_1\gamma_2$ •SIGK complex solubilized in CHAPS.	84
Figure 2-13. Optimized crystals of the G $\beta_1\gamma_2$ •SIGK complex solubilized in CHAPS.	84
Equation 3-1. The structure factor equation.....	88
Equation 3-2. The electron density equation.....	89

Figure 3-1. Graphical comparison of the conventional rotation function and the direct rotation function	94
Equation 3-3. The target of the direct rotation function	95
Equation 3-4. The correlation function in terms of structure factor intensities	95
Equation 3-5. The normalized rotation overlap function utilized in AMoRE.....	96
Equation 3-6. The calculation of structure factors in AMoRE	98
Equation 3-7. The likelihood-enhanced rotation function	99
Table 3-1. Systematic absences along the three axes of the SIGK•G $\beta_1\gamma_2$ data set used for structure determination	104
Figure 3-2. The arrangement of molecules in the P1 model.	107
Table 3-2. Refinement statistics for the SIGK•G $\beta_1\gamma_2$ complex.....	109
Figure 4-1. Structure of the peptide SIGK bound to the G $\beta_1\gamma_2$ heterodimer	118
Figure 4-2. Electrostatic potential of G $\beta_1\gamma_2$	120
Figure 4-3. Graphical representation of the contacts between G β_1 and SIGK peptide	121
Table 4-1. Description of the SIGK binding site on G β_1	122
Figure 4-4. The N terminal SIGK interface with G β_1	123
Figure 4-5. Binding of SIGK peptide to G $\beta_1\gamma_2$ mutants.....	125

Figure 4-6. The C terminal SIGK interface with $G\beta_1$	127
Figure 4-7. Analysis of the common protein interaction surface of $G\beta_1$	128
Figure 4-8. The structural changes in $G\beta$ induced by SIGK binding.	130
Figure 4-9. Superposition of the GRK2• $G\beta\gamma$ complex	131
Figure 4-10. Superposition of the phosducin• $G\beta\gamma$ complex.....	133
Figure 4-11. Superposition of the $G\alpha\beta\gamma$ heterotrimer	134
Figure 4-12. Molecular surface comparisons of $G\beta\gamma$ binding interactions	137
Figure 4-13. Structural conservation of $G\beta$ at the SIGK binding site.....	138
Table 4-2. Interaction surfaces for $G\beta_1$ binding partners.....	141
Figure 4-14. Solvent accessibility of the SIGK binding site on $G\beta$	146
Figure 4-15. Positional variation of $G\beta_1$ residues in the SIGK binding surface.....	148
Figure 5-1. MANT labeled nucleotides used in the current study.....	155
Figure 5-2. BODIPY labeled nucleotides used in the current study	158
Table 5-1. Media formulation used in selenomethionine experiments	162
Figure 5-3. Phase diagram for MAD experiments.....	164
Figure 5-4. Theoretical absorption of selenium in a seleno-methionine experiment	165
Figure 5-5. SDS-PAGE analysis of gel filtered protein from $G\alpha_{i1}$ preps.....	169

Figure 5-6. Mass spectrometry on unlabeled and selenomethionine labeled protein from $G\alpha_{i1}$ preps.....	169
Figure 5-7. Unlabeled and selenomethionine crystals from untagged $G\alpha_{i1}$ preps..	170
Table 5-2. Output of DALI server	172
Table 5-3. Description of the data sets used for structure determination	173
Figure 5-8. Structure of Ef-Tu in the GDP-bound conformation.	174
Figure 5-9. Structure of Ef-Tu in an active conformation.....	175
Figure 5-10. The structure of the $G\alpha_{i1}$ contaminant	176
Figure 5-11. The process of building nucleotide models into electron density	178
Figure 5-12. The final model for the MANT-GDP data set.	180
Figure 5-13. The final model with refined density for the BODIPY-GMPPNP data set	181
Figure 5-14. Final density surrounding the MANT-GMPPNP data set.	181
Figure 5-15. Final model of the data set containing the BODIPY-GTP γ S nucleotide	182
Figure 5-16. Superposition of the five nucleotide data sets collected	183
Figure 6-1. Schematic of GST- $G\alpha_{i1}$ and GST-mod $G\alpha_{i1}$ constructs.....	191
Figure 6-2. Purity of GST- $G\alpha_{i1}$	191

Figure 6-3. GST-modG α_{i1}	192
Figure 6-4. Purification by phenyl superose	193
Figure 6-5. Purification using HiTrapQ.....	194
Figure 6-6. Final purity of the GST-modG α_{i1} protein.....	195
Figure 6-7. GTP γ S and GDP-bound crystals of modG α_{i1}	196
Figure 6-8. Effect of ammonium sulfite on BODIPY labeled nucleotide.	198
Figure 6-9. Effect of ammonium sulfite on MANT labeled nucleotide.....	198
Figure 6-10. Effect of potassium phosphate on BODIPY labeled nucleotide	199
Figure 6-11. Effect of potassium phosphate on MANT-labeled nucleotide	200
Figure 6-12. Crystals of the GDP•G α_{i1} complex.	201
Figure 7-1. Peptides of diverse sequence bind G β using distinct residues.	205

LIST OF ABBREVIATIONS

ALS	Advanced Light Source
AmoRe	Automatic molecular replacement
APS	Advanced Photon Source
β-ARK-ct	WKKELRMTMGEEEDFFDLLASKSQSKRMDDQRVDLAG
β-ME	β -mercaptoethanol
CHAPS	3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate
CCD	Charge coupled device
CCP4	Collaborative Computing Project IV
CNS	Crystallography and NMR System
<i>e. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediamine tetra-acetic acid
FITC	Fluorescein-5-isothiocyanate
FPLC	Fast performance liquid chromatography
Gα	Heterotrimeric G protein, α subunit
Gβ	Heterotrimeric G protein, β subunit
Gγ	Heterotrimeric G protein, γ subunit
GAP	GTPase activating protein
GDI	Guanine nucleotide dissociation inhibitor
GDP	Guanosine 5'-diphosphate
GEF	Guanine nucleotide exchange factor
GIRK	G protein regulated inwardly rectifying channel
GMPPNP	Guanosine 5'-[β,γ -imido]triphosphate
GPCR	G protein coupled receptor
GRK	G protein receptor kinase
GST	Glutathione S-transferase
GTP	Guanosine 5'-triphosphate
GTPγS	Guanosine 5'-O-(3-thiotriphosphate)

HEPES	4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid
IPTG	Isopropyl β -D-thiogalactoside
kDa	kilo Dalton
LB	Luria Broth
MAD	Multi-wavelength anomalous dispersion
MAPK	Mitogen activating protein kinase
MR	Molecular replacement
Ni-NTA	Ni ²⁺ -nitro-triacetic acid
PEG	Polyethylene glycol
PH	Pleckstrin homology
PI3K	Phosphoinositol-3-kinase
PLC	Phospholipase C
PMSF	phenylmethylsulfonyl fluoride
QEHA	QEHAQEPERQYMHIGTMVEFAYALVGK
RGS	Regulators of G-protein signaling
SAD	Single wavelength anomalous dispersion
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel eletrophoresis
SIGK	SIGKARKILGYPDYD
SIRK	SIRKALNILGYPDYD
TRIS	Tris(hydroxymethyl) aminomethane

CRYSTALLOGRAPHIC TERMS AND ABBREVIATIONS

B factor	Temperature factor
d_{min}	High resolution limit (Å)
φ	Phase angle
F_c, F_{calc}	Calculated structure factor amplitude
F_o, F_{obs}	Observed structure factor amplitude
F_P	Native structure factor amplitude
F_{PH}	Calculated derivative structure factor amplitude
F_H	Calculated heavy atom structure factor amplitude
FOM	Figure of merit
I	Intensity of reflections
λ	Wavelength
P (φ)	Experimental phase probability distribution
RMSD	Root mean square deviation (Å)
σ	Sigma, standard deviation
V_m	Matthews coefficient (Å ³ /Dalton)

The crystallographic structure factor is calculated from a model using individual atomic form factors *f* for *N* atoms:

$$F(h, k, l) = \sum_{j=1}^N f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

Electron density is calculated from phased structure factors:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F(h, k, l) e^{-2\pi i(hx + ky + lz)}$$

R_{sym} is the R factor for comparing the intensities of symmetry related reflections:

$$R_{\text{sym}} = \frac{\sum_h \sum_i |I_i(h) - \langle I(h) \rangle|}{\sum_h \sum_i I_i(h)}, \text{ where } I_i(h) \text{ and } \langle I(h) \rangle \text{ are the } i\text{th and mean measurement of the intensity of reflection } h.$$

Figure of merit (FOM) is defined as $\langle \int_{\varphi} P(\varphi) \exp(i\varphi) d\varphi \rangle / \int_{\varphi} P(\varphi) d\varphi$, where *P* (φ) is the probability density function. It indicates the “sharpness” of the phase probability distribution.

R factor (**R_{work}**) is the crystallographic residual comparing a model and X-ray data:

$R_{\text{work}} = \sum_h ||F_{\text{obs}}(h)| - |F_{\text{calc}}(h)|| / \sum_h |F_{\text{obs}}(h)|$, where $F_{\text{obs}}(h)$ and $F_{\text{calc}}(h)$ are the observed and calculated structure factors, respectively.

R_{free} is the R factor value obtained for a test set of reflections consisting of a randomly selected percentage of the data.

CHAPTER ONE

Introduction

HETEROTRIMERIC G PROTEINS: BACKGROUND AND BIOLOGICAL RELEVANCE

Heterotrimeric G Protein Cycle

Heterotrimeric G proteins transduce cellular signals throughout eukaryotic cells. The heterotrimer consists of an α subunit ($G\alpha$), which encodes a GTPase, and β and γ subunits which contain no intrinsic enzymatic activity and act functionally as a single unit ($G\beta\gamma$) (Neves, Ram et al. 2002; Offermanns 2003). Heterotrimeric G protein signaling is initiated either by an activated G protein coupled receptor (GPCR) or by soluble nucleotide exchange factors (GEFs), which interact with the basal GDP-bound $G\alpha$ while it is in complex with $G\beta\gamma$ (Cismowski, Takesono et al. 2001; Marinissen and Gutkind 2001). This interaction catalyzes the release of GDP from $G\alpha$, which allows intracellular GTP to bind $G\alpha$. The binding of GTP to $G\alpha$ induces conformational changes, especially in the so-called "switch" regions, which are incompatible with $G\beta\gamma$ binding and induce dissociation of the heterotrimeric complex; signaling can then occur independently through $G\alpha$ and $G\beta\gamma$ (Sprang 1997). In the absence of any regulation, signaling ceases when the GTPase activity of $G\alpha$ hydrolyzes GTP to form GDP and inorganic phosphate. The basal form of $G\alpha$ can then reform the heterotrimeric complex with $G\beta\gamma$, for which it has high affinity in the GDP-bound form (Sprang 1997). There are many modes of regulating

heterotrimeric G protein signaling; in addition to the GEFs mentioned above that activate G protein signaling, there are guanine dissociation inhibitors (GDI) which bind to $G\alpha$ and stabilize the signaling inactive conformation. $G\beta\gamma$ is a GDI, and competition with other cellular GDIs would effectively activate signaling through unbound $G\beta\gamma$ (Sprang 1997). In addition, regulators of G protein signaling (RGS) proteins bind to $G\alpha$ and stimulate GTPase activity (Ross and Wilkie 2000), and there is evidence in at least one pathway of a $RGS \cdot G\alpha \cdot \text{effector}$ complex in which the effector further stimulates the GTPase activity of the RGS protein (He, Lu et al. 2000). See Figure 1-1 for a graphical representation of the G protein signaling cycle. Several aspects of the G protein pathway remain mysterious; examples include the molecular mechanism by which GPCRs induce nucleotide exchange (Rondard, Iiri et al. 2001); and whether subunit dissociation of $G\alpha$ from $G\beta\gamma$ is necessary for signaling to occur (Levitzki and Klein 2002).

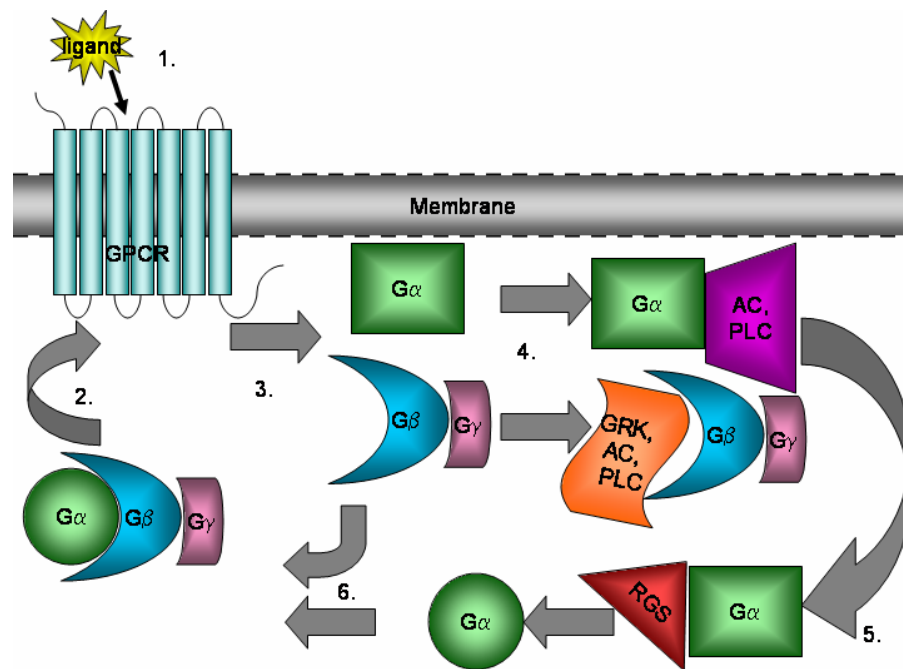


Figure 1-1. The heterotrimeric G protein cycle. A graphical representation of a basic heterotrimeric G protein cycle is shown. (1) An extracellular ligand interacts with a seven transmembrane G protein coupled receptor (GPCR), thereby inducing conformational changes which attract a resting (GDP-bound) G protein heterotrimer (2). Upon interaction with an activated receptor, $G\alpha$ undergoes a conformational change which induces GDP dissociation and GTP binding. This induces a conformational change in $G\alpha$ which is incompatible with $G\beta\gamma$ binding, and the two signaling units functionally separate (3) to interact with various effectors depending on G protein isoform and cell type (4). In the absence of regulation, this signaling would continue until the inherent GTPase activity of the $G\alpha$ subunit hydrolyzed GTP to GDP + Pi. However, in most systems G protein regulators (RGS proteins) bind to activated $G\alpha$ or $G\alpha$ -effector complex and catalyze hydrolysis (5). After hydrolysis, $G\alpha$ undergoes another conformational change which is compatible with $G\beta\gamma$ binding, and reassociation of the heterotrimer (6) ends one round of the signaling cycle. It should be noted that due to lipid post-translational modifications on $G\alpha$ and $G\gamma$, both $G\alpha$ and $G\beta\gamma$ remain associated with the membrane during the cycle, as do many of their effectors. Abbreviations: GPCR, G protein coupled receptor; AC, adenylyl cyclase; PLC, phospholipase C; RGS, regulator of G protein signaling.

G Protein α Subunits and their Relationship to $G\beta\gamma$

More than twenty $G\alpha$ genes have been described in the literature (Sprang 1997;

Offermanns 2003). $G\alpha$ subunits are separated into four classes based on

sequence: the $G\alpha_{i/o}$ ($G\alpha_{i1}$, $G\alpha_{i2}$, $G\alpha_{i3}$, $G\alpha_z$, $G\alpha_{gust}$, $G\alpha_t$), $G\alpha_q$ ($G\alpha_{11}$, $G\alpha_{14}$, $G\alpha_{15/16}$), $G\alpha_{12}$ ($G\alpha_{13}$), and $G\alpha_s$ ($G\alpha_{olf}$) families (Sprang 1997; Offermanns 2003). Many of these gene products are ubiquitously expressed ($G\alpha_s$, $G\alpha_{i2}$, $G\alpha_q$, $G\alpha_{12}$, $G\alpha_{13}$), and the $G\alpha_i$ -type family members ($G\alpha_{i1}$, $G\alpha_{i2}$, and $G\alpha_{i3}$) are the most widely expressed in cells (Offermanns 2003). Other $G\alpha$ subunits are tissue specific, residing in neurons ($G\alpha_o$) and retinal cells ($G\alpha_t$). Knockouts of several $G\alpha$ subtypes in mice are embryonic lethal, including $G\alpha_s$, $G\alpha_{13}$, $G\alpha_{i2}/G\alpha_{i3}$ and $G\alpha_{12}/G\alpha_{13}$ crossed lines; several knockouts show no obvious phenotype at all, including $G\alpha_{i1}$ and $G\alpha_{i3}$, $G\alpha_{11}$, $G\alpha_{14}$, and $G\alpha_{15}$, and $G\alpha_{12}$ (Wettschureck, Moers et al. 2004) (Table 1-1).

$G\alpha$ subunits consist of two distinct domains - a ras-like domain with structural homology to small GTPases and a helical domain that is unique to heterotrimeric α subunits (Hamm and Gilchrist 1996; Sprang 1997). Structures of GDP and activated ($GTP\gamma S$, GMPPNP, $GDP \cdot Mg^{++} \cdot AlF_4$) $G\alpha$ subunits helped to elucidate the mechanism of nucleotide binding and the molecular determinants of $G\alpha$ activation (see later sections of this chapter for a detailed structural analysis). The regions of α that undergo the largest changes between the basal and active forms of $G\alpha$ are the switch regions, corresponding to residues 175-184, 199-219, and 227-242 in the $G\alpha_{i1}$ structure (Figure 1-5). $G\alpha$ binding proteins sense the nucleotide binding state of $G\alpha$ by interacting with these switch regions; in the crystal structures of $G\alpha$ complexes (with $G\beta\gamma$, RGS4, RGS9 and $PDE\gamma$), interactions with switch residues

comprise the main interaction surface with $G\alpha$. Effectors of $G\alpha$ include adenylyl cyclase types I-IX; phosphodiesterase δ ; phospholipase-C (PLC) β isoforms 1-4; and various RhoGEF proteins (Table 1-1.) (Albert and Robillard 2002; Offermanns 2003).

Family	Subtype	Gene	Expression	Effector	Subtype	Knockout Phenotype
Gα_s	G α_s	Gnas	Ubiquitous	AC (+) , Tubulin (+), src (+)	AC I-IX	Early embryonic lethality
	G α_{olf}	Gna1	Olfactory epithelium, brain	AC (+)		Subtle; loss of fostering behavior, hyperactivity
G$\alpha_{i/o}$	G α_{i1}	Gnai1	Wide	AC (-); GIRK (+)	AC I, III, V, VI, VIII, IX	none
	G α_{i2}	Gnai2	Ubiquitous	AC (-); GIRK (+); Rap1 GAP	AC I, V, VI	Immune abnormalities and inflammation in bowel
	G α_{i3}	Gnai3	Wide	AC (-); GIRK (+); GRIN 1 and 2	AC I, V, VI	none
	G α_o	Gnao	Neuronal	AC (-); VDCC(-); GIRK(+); Tubulin (+); src (+)	N, P/Q VDCC, AC I	Various disorders, including tremors and seizures,
	G α_z	Gnaz	Neuronal, platelets	AC (-)	AC I, V, VI	Mild platelet/CNS defects
	G α_{gust}	Gnat3	Taste/brush cells	PDE (+)		Impaired bitter, sweet, and umami sensation
	G α_t	Gnat1/2	Retinal	cGMP-PDE (+)	PDE 6 γ (rod & cone)	Mild retinal degeneration
	G $\alpha_{i1}/G\alpha_{i3}$ G $\alpha_{i2}/G\alpha_{i3}$					None lethal
Gα_q	G α_q	Gnaq	Ubiquitous	PLC β (+)	PLC β 1-4	Cerebellar ataxia; impaired platelet activation
	G α_{11}	Gna11	Almost ubiquitous	PLC β (+)	PLC β 1-4	None
	G α_{14}	Gna14	Kidney, lung, spleen	PLC β (+)		None
	G α_{15}	Gna15	Hemato-poietic cells	PLC β (+)		None
	G $\alpha_q/G\alpha_{ij}$ G $\alpha_q/G\alpha_{15}$					Embryonic lethal As G α_q
Gα_{12}	G α_{12}	Gna12	Ubiquitous	RhoGEF, Btk, Gap1 ^m , cadherin; NHE-1 (+); PLD (+)		None
	G α_{13}	Gna13	Ubiquitous	P115RhoGEF, radixin, cadherin; iNOS (+)		Lethal e9.5
	G $\alpha_{12/13}$					Embryonic lethal e8.5

Table1-1. Effectors and regulators of G protein α subunits. Abbreviations: AC, adenylyl cyclase; GIRK, G protein regulated inwardly rectifying channel; GAP, G protein activating protein; GRIN, G protein regulated inducer of neurite outgrowth; VDCC, voltage dependent calcium channel; PDE, phosphodiesterase; PLC, phospholipase C; Btk, Bruton's tyrosine kinase; NHE-1, Na⁺/H⁺ exchangers; iNOS, inducible nitric oxide synthase. Culled from (Cabrera-Vera, Vanhauwe et al. 2003; Offermanns 2003; Wettschureck, Moers et al. 2004).

G Protein β : a WD-40 Repeat Protein

In comparison to the sixteen or more $G\alpha$ subunits found in mammalian cells, there are only five $G\beta$ isoforms identified to date (Simon, Strathmann et al. 1991; Albert and Robillard 2002). Four of these five isoforms share a high degree of sequence identity; $G\beta_1$ - $G\beta_4$ display 80-90% identity to each other, while the most divergent member of the family ($G\beta_5$) contains 50% identity to $G\beta_1$ - $G\beta_4$ (Figure 1-4) (Simonds and Zhang 2000). WD-40 repeat proteins belong to a superfamily of proteins that form β -propeller structures, and are often found in biological pathways as multiprotein scaffolds (Gettemans, Meerschaert et al. 2003; Roberts and Waelbroeck 2004). Proteins besides $G\beta$ containing WD-40 repeats include proteins involved in the anaphase promoting complex (Page and Hieter 1999); a protein that interacts with HIV-1 integrase (Violot, Hong et al. 2003); transcriptional regulators (Keleher, Redd et al. 1992; Hoey, Weinzierl et al. 1993), and vesicular trafficking (Pryer, Salama et al. 1993), among other biological functions. The WD-40 repeating unit consists generally of a region of variable length, followed by a core of more constant length bracketed by the sequences GH and WD (see Figure 1-2) (Neer, Schmidt et al. 1994). In $G\beta$, the WD repeat also contains an internal invariant Asp upstream of the terminal WD sequence (Wall, Coleman et al. 1995). Each WD-40 repeat in $G\beta$ folds into four antiparallel β strands; however, these four strands do not form a single structural unit. Instead, the first strand in the WD-40 repeat is actually the “D” or outer strand of a four stranded blade in $G\beta$; the next three strands in a given repeat (beginning with the GH sequence) form the inner strands of the next

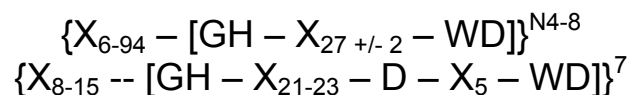


Figure 1-2. The canonical structure of the WD-40 repeat. The first line depicts the structure of a general WD-40 repeat as described by Neer et al. (Neer, Schmidt et al. 1994). The second line represents the structure of the WD-40 repeating unit found in the $G\alpha_{i1}\beta_1\gamma_2$ heterotrimer as described by Wall et al. (Wall, Coleman et al. 1995)

blade, finishing with the “D” strand of a neighboring WD-40 repeat (Figure 1-3).

Therefore, a single WD repeat would not form a well-folded structure on its own (Wall, Coleman et al. 1995). The resulting β propeller structure contains at least three types of potential interaction surface: the “top” face of the propeller formed by the loops connecting the D to A and B to C strands of the WD-40 repeat; the “bottom” face of the propeller formed by the loops connecting the A and B and C to D strands of the WD-40 repeat; and the sides of the propeller formed by the outer β strands themselves. β propeller proteins typically contain between four to eight blades; the G protein β subunit contains seven blades.

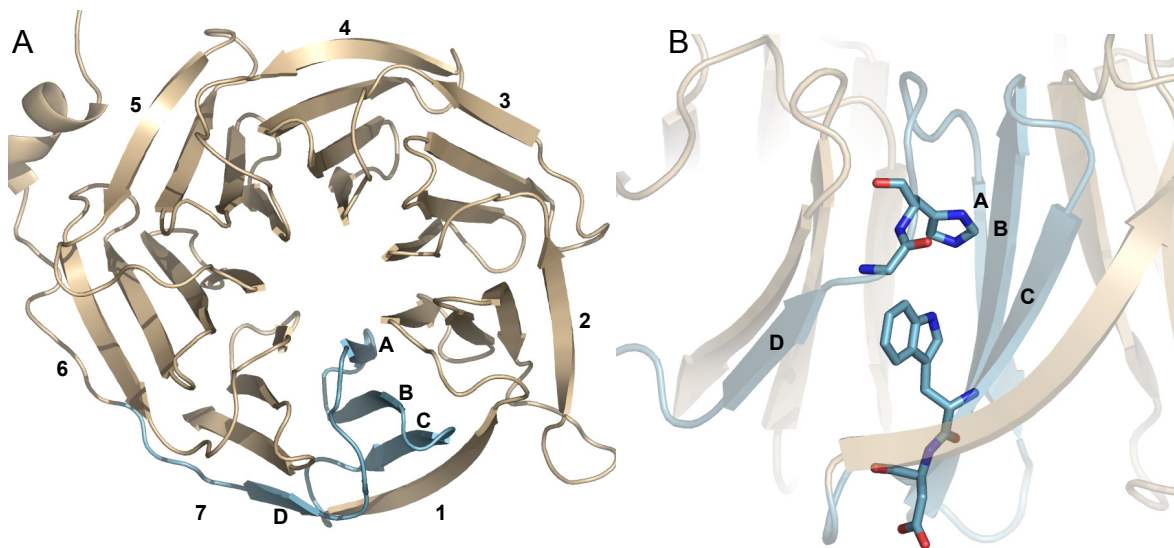


Figure 1-3. Structure of a β -propeller and a WD-40 repeat. (A) The structure of $G\beta_1$ from the $G\alpha_{i1}\beta_1\gamma_2$ structure is shown in ribbon representation. The view is looking down upon the “top” face of $G\beta$. Highlighted in blue is a single WD-40 repeat. The strands A-D are labeled. Blades are numbered along the outside of the β propeller according to Wall et al. (Wall, Coleman et al. 1995). (B) View of a single WD-40 repeat. Labels and colors are as in (A); the GH and WD signature sequences are shown in stick representation.

$G\beta_1$ - $G\beta_4$ are widely expressed, while $G\beta_5$ expression is restricted to the brain and retina (Simonds and Zhang 2000). Finally, while all other $G\beta$ subunits interact solely with $G\gamma$ subunits, $G\beta_5$ seems to functionally couple to a set of RGS proteins containing GGL (G-protein γ subunit like) domains (Snow, Krumins et al. 1998). $G\beta$ binding targets include G-protein coupled receptors, certain adenylyl cyclase subtypes, some isoforms of PLC β , G-protein inwardly rectifying potassium (GIRK) channels, calcium channels, phosphoinositol-3-kinase (PI3K), G protein receptor kinases (GRK), mitogen-activated protein kinases (MAPK), phosducin, and dyamins

(Table 1-2) (Clapham and Neer 1997; Simonds and Zhang 2000; Schwindinger and Robishaw 2001).

Family member	Gene	Expression	Effectors
G β_1	Gnb1	Generally in brain, testis, pancreas, heart, retina	PLC β 2, 3(+); AC I(-), II, IV, and VII (+); GIRK 1, 2, 4 (+); VDCC (P/Q, N, R-type) (-); GRK 2, 3 (+) ¹ ; PI3K β , γ (+); Btk (+); PKD (+); Tubulin (+); Dynamin I (+); MAPK (+) ² ;
G β_2	Gnb2		
G β_3	Gnb3		
G β_4	Gnb4		
G β_5	Gnb5		

Table 1-2. Effectors and regulators of G protein β subunits. Abbreviations are: PLC, Phospholipase C; AC, adenylyl cyclase; GIRK, G protein inwardly rectifying channel; VDCC, voltage dependent calcium channel; GRK, G protein receptor kinase; PI3K, phosphoinositide-3 kinase; MAPK; mitogen activated protein kinase. ¹Activity consists of recruitment to the membrane. ²Indirect activation. Culled from (Clapham and Neer 1997; Schwindinger and Robishaw 2001; Cabrera-Vera, Vanhauwe et al. 2003; Offermanns 2003).

hbeta1	MSELQQLRQEAQELKNQIRDARKACA	DATLSQITNN	DP	VGRIQMRTTRRT	50
mbeta1	MSELQQLRQEAQELKNQIRDARKACA	DATLSQITNN	DP	VGRIQMRTTRRT	50
rbeta1	MSELQQLRQEAQELKNQIRDARKACA	DATLSQITNN	DP	VGRIQMRTTRRT	50
hbeta2	MSELEQLRQEAQELRNQIRDARKACA	DSTLTQITAG	LD	VGRIQMRTTRRT	50
mbeta2	MSELEQLRQEAQELRNQIRDARKACA	DSTLTQITAG	LD	VGRIQMRTTRRT	50
rbeta2	MSELEQLRQEAQELRNQIRDARKACA	DSTLTQITAG	LD	VGRIQMRTTRRT	50
hbeta3	MSEMEQLRQEAQELKKQIRDARKACA	DITLAELV	SG	VGRVQMRTTRRT	50
mbeta3	MSEMEQLRQEAQELKKQIRDARKACA	DITLAELV	SG	VGRVQMRTTRRT	50
rbeta3	MSEMEQLRQEAQELKKQIRDARKACA	DITLAELV	SG	VGRVQMRTTRRT	50
hbeta4	MSELEQLRQEAQELRNQIQDARKACN	DATLVQITSN	MD	VGRIQMRTTRRT	50
mbeta4	MSELEQLRQEAQELRNQIQDARKACN	DATLVQITSN	MD	VGRIQMRTTRRT	50
rbeta4	MSELEQLRQEAQELRNQIQDARKACN	DATLVQITSN	MD	VGRIQMRTTRRT	50
hbeta5a	MAFEGLHENEFTLASLKSEAESL	SKL	EEERAKL	HDVELHQVAER	50
hbeta5L	MAFEGLHENEFTLASLKSEAESL	SKL	EEERAKL	HDVELHQVAER	50
mbeta5	MAFEGLHENEFTLASLKSEAESL	SKL	EEERAKL	HDVELHQVAER	50
rbeta5	MAFEGLHENEFTLASLKSEAESL	SKL	EEERAKL	HDVELHQVAER	50
Consensus	MSELEQLRQEAQEL-NQIRDARKAC-DATL-QITSLD-VGRIQMRTTRRT				50

hbeta1	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
mbeta1	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
rbeta1	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
hbeta2	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
mbeta2	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
rbeta2	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
hbeta3	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
mbeta3	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
rbeta3	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
hbeta4	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
mbeta4	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
rbeta4	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100
hbeta5a	FVMKTRRLTLKG	HGNKVL	CMDWCKD	KRRI	VSSSQDGKVI	100
hbeta5L	FVMKTRRLTLKG	HGNKVL	CMDWCKD	KRRI	VSSSQDGKVI	100
mbeta5	FVMKTRRLTLKG	HGNKVL	CMDWCKD	KRRI	VSSSQDGKVI	100
rbeta5	FVMKTRRLTLKG	HGNKVL	CMDWCKD	KRRI	VSSSQDGKVI	100
Consensus	LRGHLAKIYAMHWGTD	SRLLVSASQDGKLI	I	WDSYTTNKVHAI	PLRSSWV	100

hbeta1	MTCAAYAPSGNYVACGGLDNI	CSIYNLKTREGNVRVSREL	A	GHTGYLSCCR	150
mbeta1	MTCAAYAPSGNYVACGGLDNI	CSIYNLKTREGNVRVSREL	A	GHTGYLSCCR	150
rbeta1	MTCAAYAPSGNYVACGGLDNI	CSIYNLKTREGNVRVSREL	A	GHTGYLSCCR	150
hbeta2	MTCAAYAPSGNFVACGGLDNI	CSIYSLKTREGNVRVSREL	P	GHTGYLSCCR	150
mbeta2	MTCAAYAPSGNFVACGGLDNI	CSIYSLKTREGNVRVSREL	P	GHTGYLSCCR	150
rbeta2	MTCAAYAPSGNFVACGGLDNI	CSIYSLKTREGNVRVSREL	P	GHTGYLSCCR	150
hbeta3	MTCAAYAPSGNFVACGGLDNM	CSIYNLKSREGNVKVSREL	SA	GHTGYLSCCR	150
mbeta3	MTCAAYAPSGNFVACGGLDNM	CSIYNLKSREGNVKVSREL	SA	GHTGYLSCCR	150
rbeta3	MTCAAYAPSGNFVACGGLDNM	CSIYNLKSREGNVKVSREL	SA	GHTGYLSCCR	150
hbeta4	MTCAAYAPSGNYVACGGLDNI	CSIYNLKTREGNVRVSREL	P	GHTGYLSCCR	150
mbeta4	MTCAAYAPSGNYVACGGLDNI	CSIYNLKTREGNVRVSREL	P	GHTGYLSCCR	150
rbeta4	MTCAAYAPSGNYVACGGLDNI	CSIYNLKTREGNVRVSREL	P	GHTGYLSCCR	150
hbeta5a	VTMPCTWYMACAYAPSGCA	ACGGLDNKC	SVYPLTFD	KENMAAKKKSVA	150
hbeta5L	VTMPCTWYMACAYAPSGCA	ACGGLDNKC	SVYPLTFD	KENMAAKKKSVA	150
mbeta5	VTMPCTWYMACAYAPSGCA	ACGGLDNKC	SVYPLTFD	KENMAAKKKSVA	150
rbeta5	VTMPCTWYMACAYAPSGCA	ACGGLDNKC	SVYPLTFD	KENMAAKKKSVA	150
Consensus	MTCAAYAPSGNYVACGGLDNI	CSIYNLKTREGNVRVSREL	-GHTGYLSCCR		150

hbeta1	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
mbeta1	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
rbeta1	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
hbeta2	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
mbeta2	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
rbeta2	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
hbeta3	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
mbeta3	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
rbeta3	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
hbeta4	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
mbeta4	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
rbeta4	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200	
hbeta5a	R	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	F	T	G	H	T	G	D	V	M	S	L	S	L	A	P	D	T	R	L	F	V	200
hbeta5b	M	H	T	N	L	S	A	C	S	F	T	N	S	D	M	Q	L	T	A	S	G	D	T	C	A	L	M	D	V	E	S	G	R	L	L	S	F	H	G	H	A	D	V	L	C	200					
hbeta5L	M	A	C	A	Y	A	P	S	G	C	A	C	G	L	D	N	K	S	Y	P	L	D	K	N	E	N	M	A	A	K	K	S	V	A	M	H	T	N	L	S	A	200									
mbeta5	M	H	T	N	L	S	A	C	S	F	T	N	S	D	M	Q	L	T	A	S	G	D	T	C	A	L	M	D	V	E	S	G	R	L	L	S	F	H	G	H	A	D	V	L	C	200					
rbeta5	M	H	T	N	L	S	A	C	S	F	T	N	S	D	M	Q	G	C	K	A	M	Y	W	D	M	R	S	G	Q	C	V	A	F	E	L	F	H	E	S	D	V	K	E	Y	200						
Consensus	F	L	D	D	N	Q	I	V	T	S	S	G	D	T	T	C	A	L	W	D	I	E	T	G	Q	Q	T	T	-	F	-	G	H	T	G	D	V	M	S	L	S	L	A	P	D	-	R	L	F	V	200

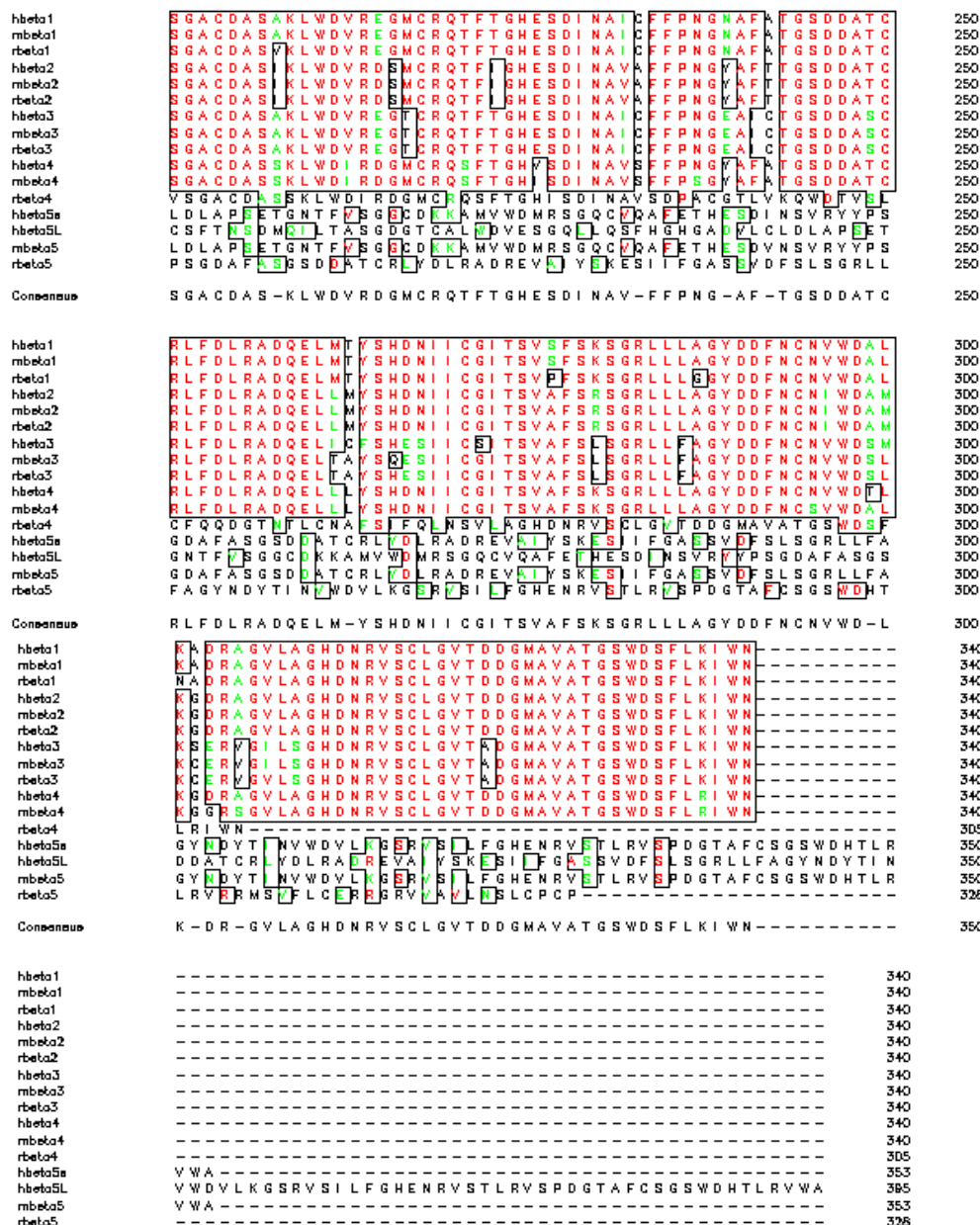


Figure 1-4. Sequence alignment of representative mammalian G β isoforms. Shown are G β isoforms 1-5 for human, mouse, and rat sequences aligned using CLUSTALW. Identical residues are shown in red, conservative substitutions in green, and non-conserved residues in black. A consensus sequence is shown at the bottom of each frame.

G Protein γ : Formation of an Obligate Heterodimer with $G\beta\gamma$ and Isoform Specificity

There are at least 14 $G\gamma$ isoforms identified in mammals to date (Albert and Robillard 2002), which show greater amino acid diversity than seen in the $G\beta$ subunit (Clapham and Neer 1997; Robishaw and Berlot 2004). Although it is possible to express and purify $G\beta$ and $G\gamma$ from insect cells separately, they form a functional unit only when coexpressed in the cell (Iniguez-Lluhi, Simon et al. 1992; Ueda, Iniguez-Lluhi et al. 1994). $G\beta$ has never been purified from native sources without a bound $G\gamma$ (or GGL protein), indicating that the $G\beta\gamma$ heterodimer is the functional unit in cells.

Although much of the literature treats all $G\beta\gamma$ heterodimers as interchangeable, the large number of different combinations of $G\beta$ and $G\gamma$ subunits most likely plays a role in specificity in intact signaling pathways. Most combinations of $G\beta\gamma$ isoforms can form functional dimers *in vitro* that can then bind to $G\alpha$ subunits (Iniguez-Lluhi, Simon et al. 1992; Ueda, Iniguez-Lluhi et al. 1994; Hou, Azpiazu et al. 2000; Hou, Chang et al. 2001). In addition, many $G\beta\gamma$ isoforms can bind to and modulate effector function, at least in the cases of adenylyl cyclase (Iniguez-Lluhi, Simon et al. 1992; Ueda, Iniguez-Lluhi et al. 1994; Hou, Chang et al. 2001), phospholipase C $\beta 2$ and $\beta 3$ (Ueda, Iniguez-Lluhi et al. 1994; Hou, Azpiazu et al. 2000; Hou, Chang et al. 2001) and GIRK channels (Mirshahi, Robillard et al. 2002). However, current literature suggests that the importance of $G\beta\gamma$ isoform specificity may manifest itself

at the level of the GPCR; for instance, multiple reports have shown that the isoform type of $G\beta$ and $G\gamma$ subunits have differential effects on Ca^{++} channel regulation and GPCR signaling ((Birnbaumer 1992; Albert and Robillard 2002). It has also been shown that receptors can discriminate between different heterotrimer compositions; for instance, M2 muscarinic receptor differentially activates $G\alpha_o\cdot\beta_1\gamma_7$ or $G\alpha_o\cdot\beta_4\gamma_2$ heterotrimers over $G\alpha_o\cdot\beta_1\gamma_5$ or $G\alpha_o\cdot\beta_1\gamma_2$ heterotrimers (Hou, Azpiazu et al. 2000; Hou, Chang et al. 2001). In another study, this time with the α_2A adrenergic receptor, different combinations of $G\beta_1$ or $G\beta_3$ with $G\gamma_1$, $G\gamma_2$, $G\gamma_3$, $G\gamma_4$, $G\gamma_5$, $G\gamma_7$, $G\gamma_{10}$, or $G\gamma_{11}$ led to different coupling efficiencies through the $G\alpha_i$ heterotrimer (Richardson and Robishaw 1999). Finally, a recent knockout of the $G\gamma_7$ isoform in mice showed that the D1 dopamine receptor signaling pathway was disrupted, as were levels of $G\alpha_{olf}$, indicating that at the organism level isoform specificity is crucial for proper function (Schwindinger, Betz et al. 2003). Recent studies have attempted to use labeling systems to identify the native $G\beta\gamma$ combinations found in intact cell lines (Hynes, Tang et al. 2004). However, there is little information on the biological effect of having different populations of $G\beta\gamma$ heterodimers in living cells, and the regulatory process mediating heterodimer formation is unknown. Temporal and spatial regulation in the cell at the level of transcription may be one level of regulation (Robishaw and Berlot 2004).

G $\beta\gamma$ Effector Biology

The most well-studied isoforms combinations of G $\beta\gamma$ are G $\beta_{1\gamma_1}$, found in the transducin pathway, and G $\beta_{1\gamma_2}$, which regulates adenylyl cyclase through coupling to G α_s or G α_i . Studies have tried to delineate the regions on G β that mediate effects on adenylyl cyclase, phospholipase C β isoforms, protein kinase C, potassium and calcium channels; and indirectly on GRK and MAPK signaling (Tang and Gilman 1991; Koch, Inglese et al. 1993; Inglese, Luttrell et al. 1994; Yan and Gautam 1996; Ford, Skiba et al. 1998; Krapivinsky, Kennedy et al. 1998; Li, Sternweis et al. 1998; Scott, Huang et al. 2001; Albert and Robillard 2002; Doering, Kisilevsky et al. 2004). The mutations that effect G β signaling through some of these effectors, as well as residues shown to interact with G β through analysis of the crystal structures of complexes are shown in Table 1-3. As shown in Figure 1-3, the structural features of the G $\beta\gamma$ complex can be described on a general level as consisting of the “top”, “bottom”, and “sides” of the G β torus. The “bottom” face of G β is the one against which G γ packs; the “top” face is the face opposite from the G γ binding site; and the “sides” of G γ are the β sheets that make up the blades of the β propeller. Almost all G $\beta\gamma$ -binding proteins use the “top” face of G β to interact with the G $\beta\gamma$ heterodimer; in addition, many G β targets also use residues along the side of G β blades (Figure 1-3, Table 1-3). This dual binding mode may confer specificity upon the G $\beta\gamma$ target; another possibility is that multiple binding sites confer additional energy nodes to the G $\beta\gamma$ interaction that allows for discrimination among G $\beta\gamma$ binding partners. Finally,

the binding of multiple $G\beta\gamma$ targets in signaling pathways can be thought of as a dynamic process involving multiple effectors and (possibly) regulators in a given system. Perhaps the transient binding of multiple protein targets to multiple blades of $G\beta\gamma$ allows for dynamic signaling regimes to be sampled during the lifetime of a cellular signaling event; association with a blade leading to preferential association with the top face of $G\beta$, then leading to productive signal transduction while other targets associate with other blades.

Gα _{i1}	Phosducin	GRK2	PLC β	AC	GIRK	Ca ⁺⁺
42	42					
44	44					
46	46					
47	47					
52						
53						
55		55		55	55	55
57	57	57		57		57
59	59	59			59	
75	75	75				
76		76				
78		78		78	78	78
80			80		80	
88						
89			89	89	89	
90						
91						
92						
		95				
		96				
	98	98				
99	99	99	99	99	99	
101	101	101	101	101		101
117	117	117	117	117		117
119			119	119		119
132						
143			143			143
144						
145	145	145				
	162					
182						
186	186		186	186		186
188	188	188				
204	204	204				
228	228		228	228	228	228
230	230					
246	246	246	246	246		
	274					
	290	290				
	292					
	304					
	310					
	311					
	314	314				
332	332	332	332	332		332

Table 1-3. List of Gβ binding residues derived from mutational and structural studies. Each number corresponds to a Gβ residue. Structural analysis was done using information from (Gaudet, Bohm et al. 1996; Wall, Posner et al. 1998; Lodowski, Pitcher et al. 2003). PLC β, mutational analysis of the PLC β2/3•Gβ₁γ₂ complex (Ford, Skiba et al. 1998; Li, Sternweis et al. 1998); AC, mutational analysis of the adenylyl cyclase type I/II•Gβ₁γ₂ complex (Ford, Skiba et al. 1998; Li, Sternweis et al. 1998); GIRK, mutational analysis of Gβ₁γ₂ interaction with the GIRK1/4 channel (Ford, Skiba et al. 1998); Ca⁺⁺, mutational analysis of Gβ₁γ₂ interaction with the N and P/Q type calcium channels (Ford, Skiba et al. 1998; Agler, Evans et al. 2003). Residues on the “top” face of Gβ are in blue boxes; all other residues map to blades along the side of the Gβ torus.

Peptides as G $\beta\gamma$ Targets: Theory and Biological Impact Studies

There are at least two methods of using peptides to probe protein•protein interaction surfaces. One method is simply to identify a known protein target of the protein of interest, and synthesize peptides with sequences derived from that target for use in competition studies. Not only does this enable researchers to isolate small fragments from sometimes unwieldy proteins for use in their studies, but in some instances the biological function of the binding can be recapitulated by the peptide. Another method is to screen peptide libraries of random sequences for their ability to bind to the protein of interest. Although the library-derived peptides have no inherent biological relevance, the resulting protein•peptide complexes can often be used to probe the binding sites, biochemistry, or biological activity of the protein of interest. This method as a means of probing protein-protein interaction “hot spots” has been reviewed extensively in the literature (Clackson and Wells 1994; Ma, Wolfson et al. 2001; DeLano 2002). There are many examples of an untrained (naïve) peptide library being used to delineate a binding surface that is biologically relevant (DeLano, Ultsch et al. 2000; Harris, Hillier et al. 2001; Scott, Huang et al. 2001; Skelton, Chen et al. 2001). It may be more efficient to describe binding sites this way in many pathways, since the protein binding targets for a given protein may take years to discover and verify.

G protein $\beta\gamma$ subunits have been studied using both of the methods described above. Peptides derived from adenylyl cyclase II (QEHA: residues 956-982,

QEHAQEPERQYMHIGTMVEFAYALVGK), GRK (β -ARK ct: residues 643-670, WKKELRMTMGEEFFDLLASKSQSKRMDDQRVDLAG), PLC β 2 (N20K: residues 564-583, NRSYVISSFTELKAYDLLSK and E20K: residues 574-593, ELKAYDLLSKASVQFVDYNK) GIRK1 and GIRK4 (various, including GIRK1 residues 364-383, MLLMSSPLIAPAITNSKERHNS), and phosducin (residues 215-232, EGMCRQTFTGHESDINAI have all been used to elucidate $G\beta\gamma$ biological activity in relevant pathways (Koch, Inglese et al. 1993; Weng, Li et al. 1996; Bluml, Schnepf et al. 1997; Krapivinsky, Kennedy et al. 1998; Sankaran, Osterhout et al. 1998). In general, these peptides have much lower affinity for $G\beta\gamma$ (μM – mM) than the proteins they are derived from, and present solubility problems that make them difficult to study in crystallographic concentrations. In 2001, Scott et al. used a random library approach to isolate peptides that bound to $G\beta\gamma$ (Scott, Huang et al. 2001). Sixteen different libraries were tested; the libraries contained peptides from six to thirty amino acids long, some of which contained internal disulfide linkages to impose structural constraints (Scott, Huang et al. 2001). Each library contained diversity of 1×10^8 – 1×10^9 individual clones (Scott, Huang et al. 2001). Biotinylated $\beta\gamma$ protein was immobilized onto plates containing streptavidin, phage displaying peptide were mixed with the immobilized $\beta\gamma$ subunits and repeatedly washed with detergent, bovine serum albumin, and salt to reduce nonspecific binding. Binding of phage was then detected by using ELISA with phage-specific antibody and a colorimetric assay (Scott, Huang et al. 2001). Four classes of peptides (based on sequence) were identified; all four classes competed for binding on $G\beta\gamma$, indicating

that they were all probing the same surface on $G\beta\gamma$ (see Chapter Two for more detailed information on peptide sequences). One of these groups included a linear peptide (the “SIRK” peptide) with the sequence SIRKALNILGYPDYD. The SIRK peptide inhibited PLC $\beta 2$ activation by $G\beta_{1\gamma 2}$ subunits with an IC_{50} of 5 μM and blocked activation of PI3K. In contrast, the SIRK peptide had little or no effect on $G\beta_{1\gamma 2}$ regulation of type I adenylyl cyclase or voltage-gated N-type Ca^{++} channel activity (Scott, Huang et al. 2001). Subsequent experiments have shown that not only does the SIRK peptide block heterotrimer formation, but it also displaces $G\alpha_{i1}$ from a $G\beta_{1\gamma 2} \cdot G\alpha_{i1}$ complex in the absence of $G\alpha_{i1}$ activation and activates G protein-dependent ERK1 and ERK2 pathways in intact cells (Ghosh, Peterson et al. 2003; Goubaeva, Ghosh et al. 2003). A version of the SIRK peptide is discussed below in the description of the dissertation research.

HETEROTRIMERIC G PROTEINS: STRUCTURAL STUDIES

Crystal Structures of Active and Inactive Forms of $G\alpha$

The crystal structures of inactive (GDP) and active ($GDP \cdot Mg^{++} \cdot AlF_4^-$, $GTP\gamma S \cdot Mg^{++}$, $GMPPNP \cdot Mg^{++}$) provide insight into the molecular mechanism of signaling employed by $G\alpha$ subunits (Coleman, Berghuis et al. 1994; Mixon, Lee et al. 1995; Sprang 1997; Coleman and Sprang 1998; Coleman and Sprang 1999; Lambright, Noel et al. 1994). The Ras homology domain contains structural motifs involved in nucleotide binding, recognition, and GTPase activity, the α -helical domain is unique to heterotrimeric G proteins (Hamm and Gilchrist 1996; Sprang 1997). The nucleotide

binds in a cleft between the two domains; the regions of $G\alpha$ which change conformation are called the “switch” regions, numbered I-IV. The GDP-bound form of the $G\alpha_{i1}$ protein (1GDD) has disordered switch II and switch III regions and the N terminus forms a microdomain with the C terminus which in turn interacts with a neighboring $G\alpha$ subunit (Mixon, Lee et al. 1995; Sprang 1997). This interaction is organized by a sulfate ion from the crystallization condition (Sprang 1997).

However, it should be noted that this microdomain is not observed in the structure of $G\alpha_t$ bound to GDP, and in that structure all of the switch regions are ordered (Figure 1-5) (Lambright, Noel et al. 1994). In a representative activated structure (1GIA) of $G\alpha_{i1}$ bound to $GTP\gamma S \cdot Mg^{++}$, all three switch regions are ordered and have changed conformation (Coleman, Berghuis et al. 1994). In addition, the N and C termini are disordered in the 1GIA structure. Other than these two differences, the GDP-bound and the $GTP\gamma S \cdot Mg^{++}$ $G\alpha$ structures have quite similar global conformations (Figure 1-5), indicating that the changes in the switch regions upon GTP binding are the structures accessed by effectors and regulators of $G\alpha$ (Figure 1-6).

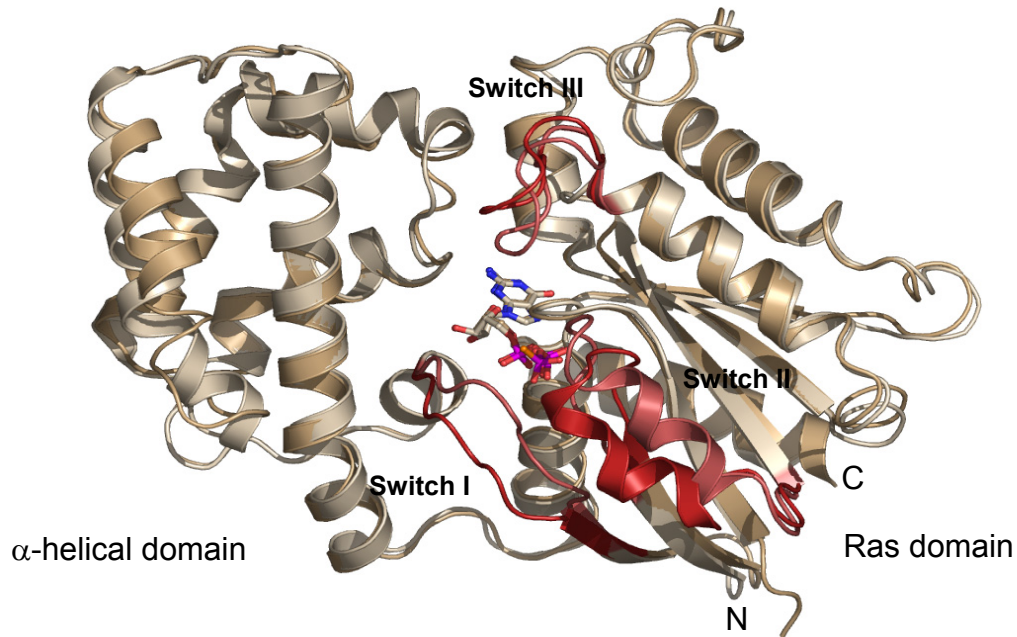


Figure 1-5. Comparison of the basal and active forms of $G\alpha$. The GDP-bound structure of $G\alpha_t$ (1TAG) (dark brown) is shown superimposed upon the structure of $G\alpha_i$ bound to $GTP_{\gamma}S \cdot Mg^{++}$ (1GIA) (light brown) (Coleman, Berghuis et al. 1994; Lambright, Noel et al. 1994). The α -helical domain and Ras homology domains are labeled. The proteins are shown in ribbon representation; the switch regions are colored in red (dark for active form, light for inactive form) and labeled. The nucleotides are shown in stick representation.

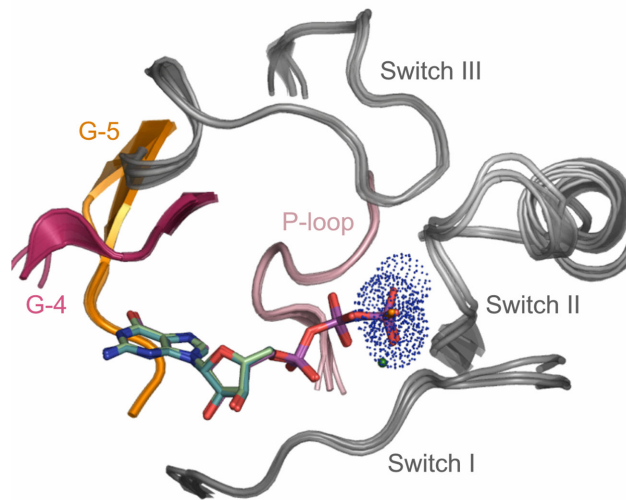


Figure 1-6. Nucleotide binding pocket interactions in $G\alpha_{i1}$. Structures of $G\alpha_{i1}$ bound to $GTP\gamma S \cdot Mg^{++}$ (1GIA), $GDP \cdot Mg^{++} \cdot AlF_4^-$ (1GFI), and $GMPPNP \cdot Mg^{++}$ (1CIP) are shown superimposed (Coleman, Berghuis et al. 1994; Coleman and Sprang 1999). Features of the binding pocket contributed in each structure are shown in ribbon representation and labeled. The dot representation is used to show the location of aluminum fluoride in the 1GFI structure, which overlaps nicely with the other activating nucleotides in the γ phosphate location. A green ball is used to represent the catalytic Mg^{++} ion found in all three structures. The G-4 and G-5 regions contain conserved sequences which confer guanine nucleotide specificity upon the protein (Sprang 1997); the P-loop is a conserved sequence found in many nucleotide-binding proteins; and the switch regions I, II, and III are the regions found in heterotrimeric G protein α subunits that undergo the largest conformational changes when bound to GDP versus GTP.

The Crystal Structure of Unbound $G\beta\gamma$

The crystal structure of unbound $G\beta_1\gamma_1$ was solved in 1996 (Sondek, Bohm et al. 1996). The features of the $G\beta_1\gamma_1$ structure are essentially the same as those of the $G\beta_1\gamma_2$ structure that is found as part of the $G\alpha_{i1} \cdot G\beta_1\gamma_2$ heterotrimer and of $G\beta_1\gamma_1$ bound to $G\alpha_t$ (Wall, Coleman et al. 1995; Lambright, Sondek et al. 1996; Wall, Posner et al. 1998). The $G\beta$ molecule is comprised of two domains: an N terminal

helix that interacts extensively with the $G\gamma$ subunit and a C terminal domain that is the β -propeller fold (Figure 1-7). The $G\beta$ propeller is composed of seven blades as discussed earlier in this chapter; the top and bottom of each blade are connected by loops of various length and diverse sequence.

The $G\gamma$ subunit is made up of two helical segments joined by a loop with no secondary structural elements. The N terminal helical segment engages the N terminal helix of $G\beta$ in a coiled-coil interaction; the C terminal helical segment packs against blades five and six of $G\beta$.

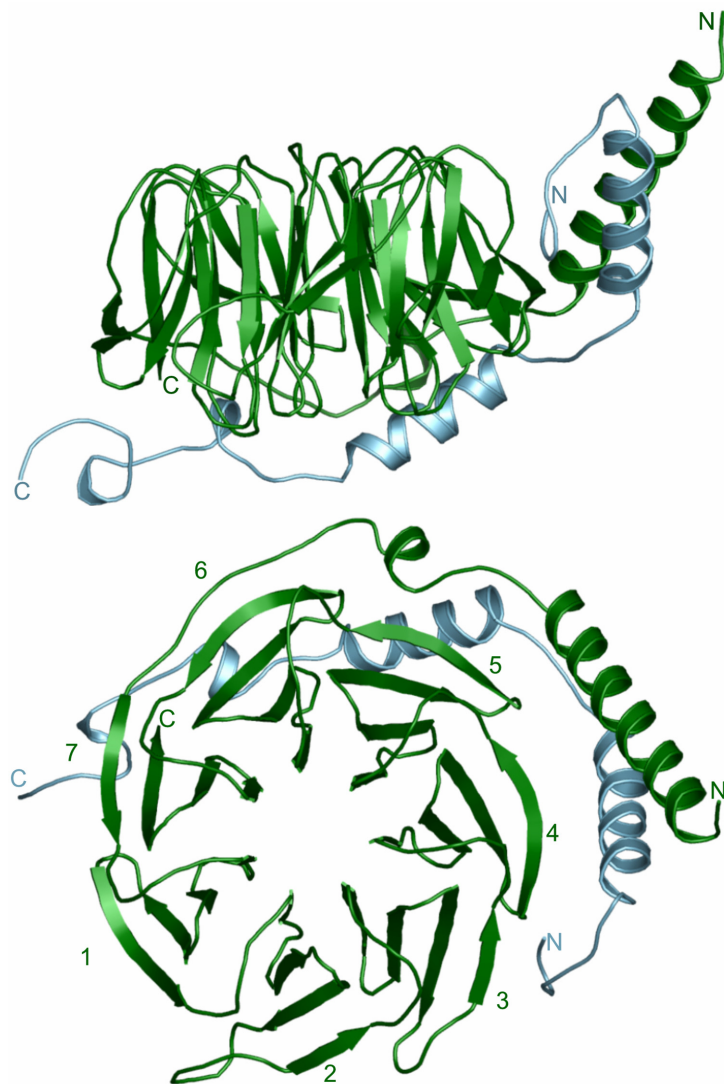


Figure 1-7. Structure of the unbound form of Gβ₁γ₁. The general features of the Gβ₁γ₁ heterodimer are shown (Sondek, Bohm et al. 1996). Gβ₁ is shown in green, Gγ₁ in blue. In the upper panel, a side view is shown; the lower panel is rotated by 90° about the x axis. N and C termini are labeled in each panel; in the lower panel, the blades of Gβ are labeled according to the convention of Wall et al. (Wall, Coleman et al. 1995; Wall, Posner et al. 1998).

The Crystal Structure of the $G\alpha\cdot\beta\gamma$ Heterotrimer

The heterotrimeric structures of $G\alpha_{i1}\cdot G\beta_{1\gamma 2}$ and $G\alpha_t\cdot G\beta_{1\gamma 1}$ delineate two interaction surfaces between $G\alpha$ and $G\beta\gamma$ (Wall, Coleman et al. 1995; Lambright, Sondek et al. 1996; Wall, Posner et al. 1998). The first is the interaction between the switch I and II region of $G\alpha$ and the “top” face of $G\beta$ (Figure 1-8). The second is an extensive interface between the N terminus of $G\alpha$ and blade 1 of $G\beta$. The switch II interaction with $G\beta$ is thought to be the molecular mechanism by which $G\beta\gamma$ acts as a guanine dissociation inhibitor, by stabilizing the GDP-bound conformation of $G\alpha$ and competing with effector contacts in the switch regions of $G\alpha$. All three switch regions of $G\alpha$ bound to $G\beta\gamma$ are remodeled relative to their conformations in the GDP or $GTP\gamma S\cdot Mg^{++}$ structures (Sprang 1997) (Figure 1-9). The N terminal interaction between $G\alpha$ and $G\beta\gamma$ is not known to play any specific role in signaling, but it is important for heterotrimer formation; deletion of the N terminus of $G\alpha$ abrogates binding to $G\beta\gamma$, and buries about 900 Å² of solvent-accessible surface (Sprang 1997).

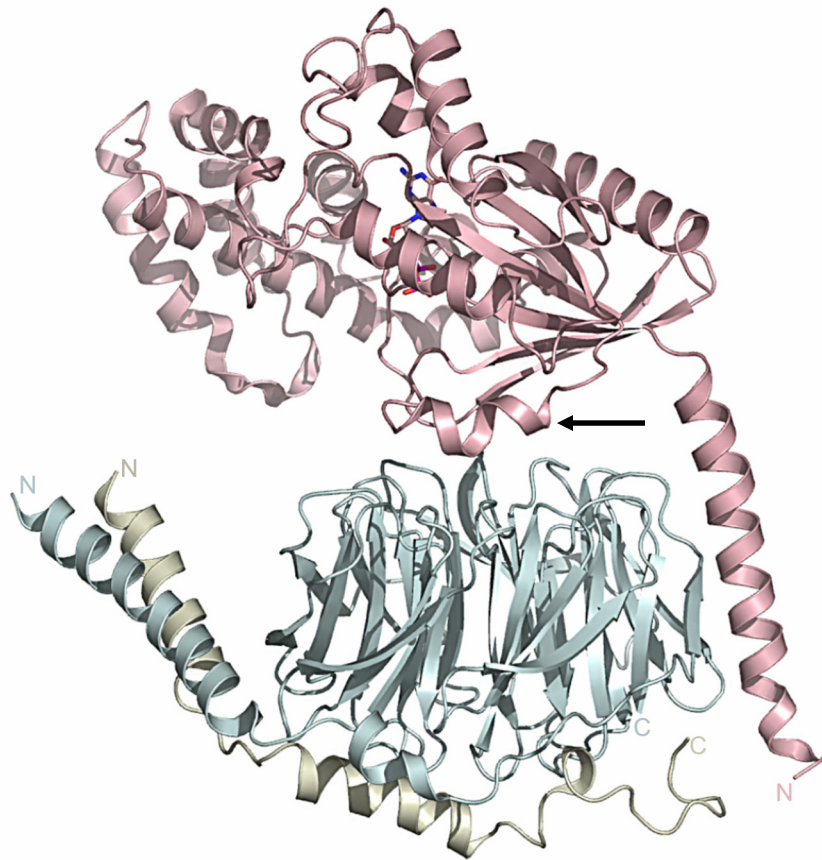


Figure 1-8. Structure of the $G_{\alpha i1} \cdot G_{\beta 1} \cdot G_{\gamma 2}$ heterotrimer. The structure of the $G_{\alpha i1} \cdot G_{\beta 1} \cdot G_{\gamma 2}$ as determined by Wall et al. (Wall, Coleman et al. 1995; Wall, Posner et al. 1998) is shown in ribbon representation. $G_{\alpha i1}$ is shown in pink; $G_{\beta 1}$ in cyan; $G_{\gamma 2}$ in khaki. The N and C terminus of each molecule is labeled; the GDP nucleotide is shown in stick representation. The view is side-on with respect to the $G_{\beta 1} \gamma 2$ heterodimer, and an arrow is pointing at the switch II helix of G_{α} where it interacts with the “top” face of G_{β} . The N terminal interaction between G_{α} and $G_{\beta 1} \gamma 2$ is clearly visible in this orientation along the right side of the figure.

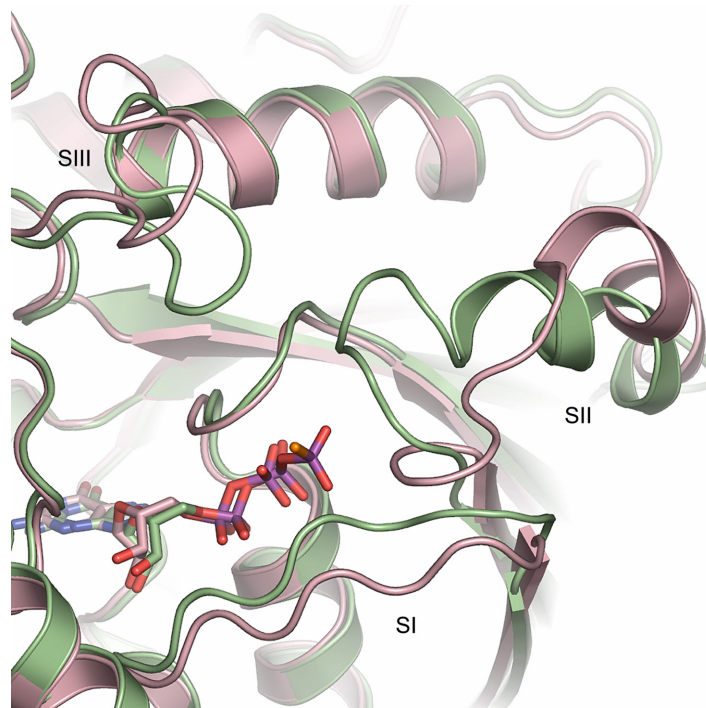


Figure 1-9. Differences in the nucleotide binding pocket between $G\alpha$ in the heterotrimer and GTP-activated $G\alpha$. The $G\alpha_{i1}$ subunits from the GDP form found in the heterotrimer structure (1GP2, in pink) and in the $GTP\gamma S$ bound structure (1GIA) are superimposed (Coleman, Berghuis et al. 1994; Wall, Coleman et al. 1995; Wall, Posner et al. 1998). Switch regions I, II, and III are shown in ribbon representation; the nucleotides are shown in stick representation. The main contact surface donated by $G\alpha$ to complex with $G\beta$ is the switch II region; the switch I and III surfaces interact with effectors and regulators and reflect the nature of the nucleotide bound by $G\alpha$.

Crystal Structures of G Proteins in Complex with Effectors and Regulators

$G\alpha$ Bound to Protein Targets

The predictions based on the $G\alpha \cdot GDP$, $G\alpha \cdot GTP\gamma S \cdot Mg^{++}$, and $G\alpha \cdot \beta\gamma$ structures included the hypothesis that interactions with the switch regions of $G\alpha$ would be crucial for interaction between $G\alpha$ and its effectors or regulators. This hypothesis has since been borne out by the structures of several $G\alpha$ complexes with regulators,

and one effector: the co-crystal structure of $G\alpha_s$ bound to the catalytic domains of adenylyl cyclase (Tesmer, Sunahara et al. 1997). There is also an interesting structure of a $G\alpha$ subunit bound to a regulator (RGS9) and effector ($PDE\gamma$) simultaneously (Slep, Kercher et al. 2001). The structure of $G\alpha$ bound to the regulator RGS4 is representative of the switch interactions exploited by $G\alpha$ binding proteins and is the first structure to visualize these interactions (Tesmer, Berman et al. 1997) (Figure 1-10).

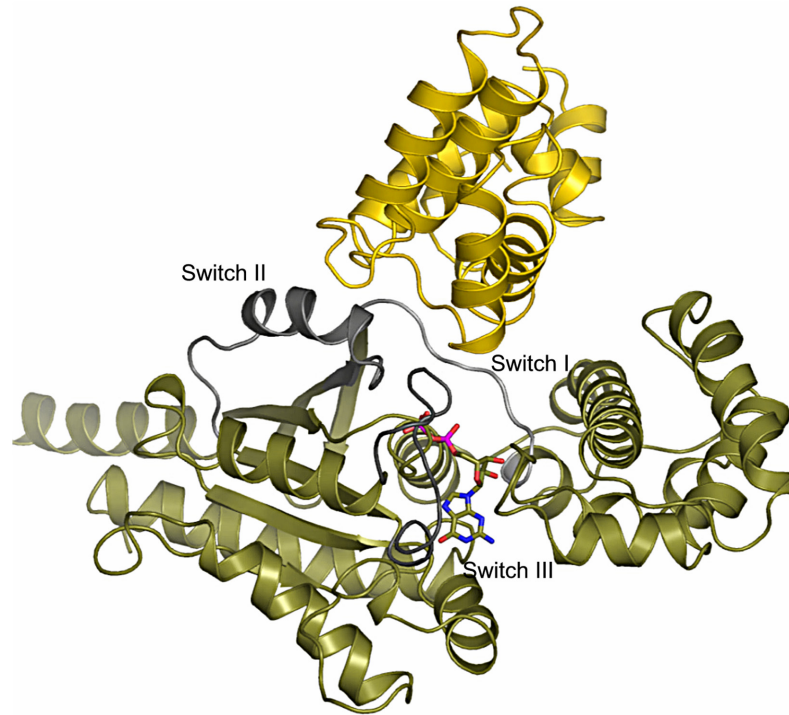


Figure 1-10. $G\alpha_{i1}$ bound to RGS4 shows the interaction of switch regions with signal transduction regulators. The structure of $G\alpha_{i1}$ bound to RGS4 (1AGR) is shown in ribbon representation (Tesmer, Berman et al. 1997). $G\alpha$ is in olive, with the switch regions colored grey and labeled. RGS4 is shown in gold. The changes in the switch regions induced by activation of the $G\alpha$ subunit are read by effectors and regulators, which interact extensively with these switches only when they are in the correct conformational state.

RGS proteins bind to the transition state analog bound form of $G\alpha$ ($GDP \cdot Mg^{++} \cdot AlF_4^-$) with the highest affinity (Berman, Kozasa et al. 1996), and although they bind the switch regions of $G\alpha$, they do not contribute any catalytic residues to the active site of $G\alpha$ (Bohm, Gaudet et al. 1997; Tesmer, Berman et al. 1997). Enhancement of catalysis seems to be strictly a consequence of stabilizing the conformation of the switch regions of $G\alpha$ in order to lower the energetic cost of forming the transition state complex (Tesmer, Berman et al. 1997).

G $\beta\gamma$ Bound to Protein Targets

Three structures of $G\beta\gamma$ bound to protein targets have been determined: the $G\alpha\beta\gamma$ heterotrimers, the complex of GRK2 with $G_1\beta_2$, and the complex of phosducin with $G\beta_1\gamma_1$ (Wall, Coleman et al. 1995; Gaudet, Bohm et al. 1996; Loew, Ho et al. 1998; Wall, Posner et al. 1998; Lodowski, Pitcher et al. 2003). The heterotrimer was the first example of a $G\beta\gamma$ •regulator complex; although $G\beta\gamma$ is usually thought of as a $G\alpha$ regulator (a GDI), $G\alpha$ also serves to remove $G\beta\gamma$ from its signaling partners and can be thought of as an inhibitor of $G\beta\gamma$ signaling. GRK2 is a multifunctional signaling protein; it contains an RGS homology domain to bind $G\alpha$ subunits, a protein kinase domain that phosphorylates GPCRs to desensitize them, and a pleckstrin homology (PH) domain that interacts with $G\beta\gamma$ (Figure 1-11) (Lodowski, Pitcher et al. 2003). Although the GRK2 protein is a complex modular structure, the interaction with the $G\beta\gamma$ subunit is actually straightforward; four regions within the PH domain of GRK2

interact exclusively with the top face of G β , utilizing many of the same residues accessed by the switch I and switch II regions of G α (Lodowski, Pitcher et al. 2003). There are no large conformational changes in G $\beta\gamma$ upon GRK2 binding, compared with either the unbound G $\beta\gamma$ structure or that of the G $\alpha\beta\gamma$ heterotrimer.

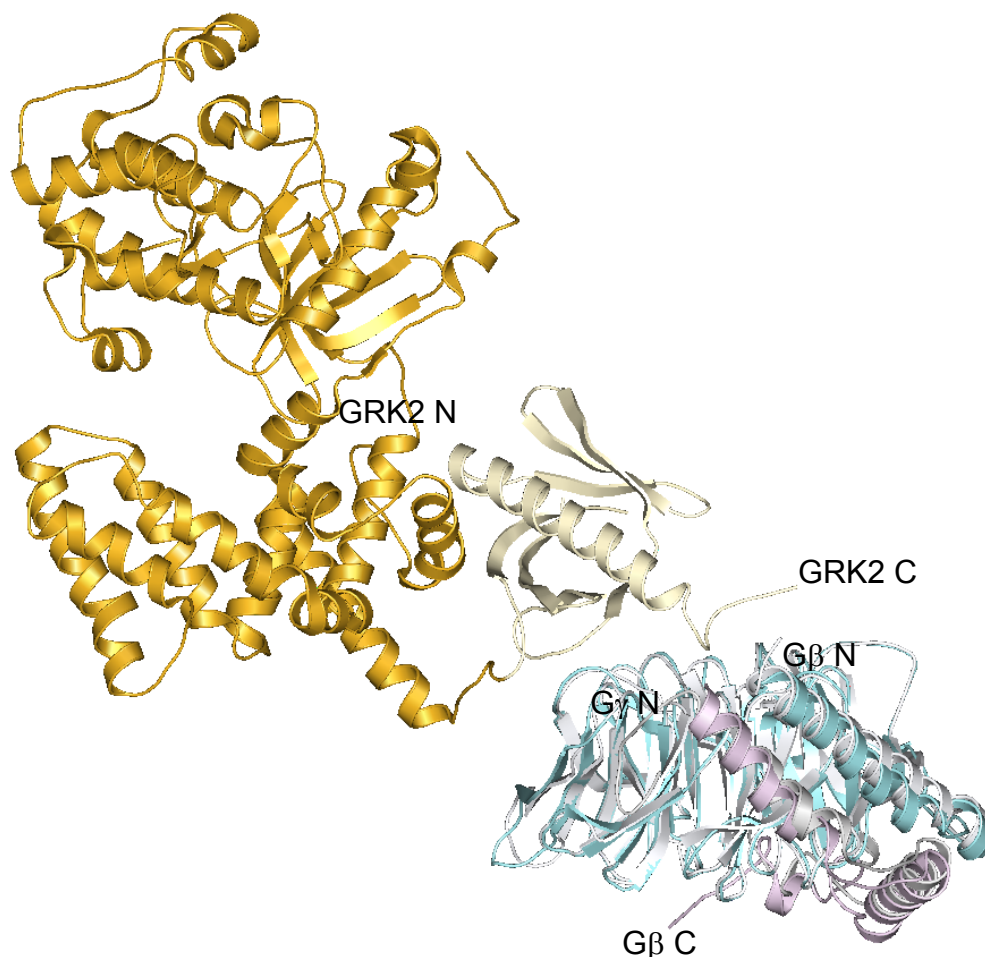


Figure 1-11. Structure of the GRK2•G $\beta_1\gamma_2$ complex. The structure of the GRK2•G $\beta_1\gamma_2$ complex (1OMW) is shown superimposed upon the unbound structure of G $\beta_1\gamma_1$ (1TBG) (Sondek, Bohm et al. 1996; Lodowski, Pitcher et al. 2003). GRK2 is in orange, with the C terminal PH domain that interacts with G β shown in ivory. G β_1 from 1OMW is in teal, G γ_2 from 1OMW is purple, and G $\beta_1\gamma_1$ is in grey. The superposition of the two G $\beta\gamma$ structures shows that G $\beta\gamma$ does not undergo large conformational changes when bound to an effector, which is also indicated by the RMSD between the two structures (0.83 Å for C α of residues G β 40-340).

Phosducin regulates $G\beta_1\gamma_1$ in the transducin signaling pathway by binding $G\beta\gamma$ and blocking reassociation with $G\alpha_t$ (Gaudet, Bohm et al. 1996; Loew, Ho et al. 1998). This inhibits signaling through rhodopsin by removing the pool of $G\alpha_t\beta_1\gamma_1$ heterotrimers that can associate productively with the receptor (Gaudet, Bohm et al. 1996; Loew, Ho et al. 1998). Although part of the phosducin structure is disordered, the interaction surfaces with $G\beta\gamma$ are well defined and are comprised of residues from both the N and C terminal domains of phosducin. The main interaction surfaces on $G\beta\gamma$ are once again the “top” face of $G\beta$, which contributes many of the same residues to phosducin as to the $G\alpha\beta\gamma$ heterotrimer and GRK2 interfaces; and interactions with the C terminal phosducin domain, which pack against blades 1 and 7 from $G\beta$ (Figure 1-12).

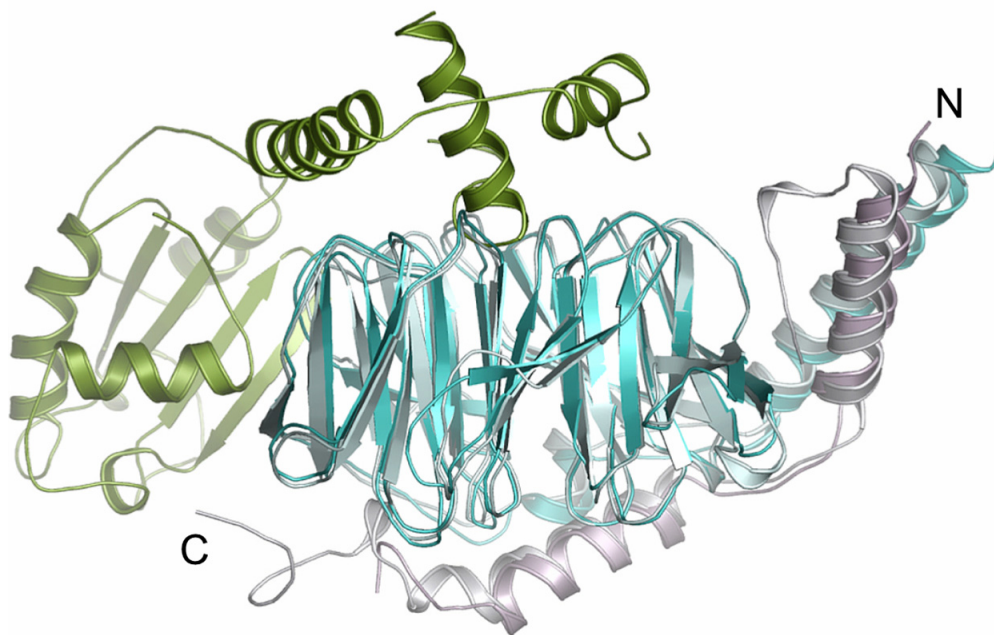


Figure 1-12. Structure of the phosducin• $G\beta_1\gamma_1$ complex. As in Figure 1-11, the structure of the phosducin• $G\beta_1\gamma_1$ complex (2TRC) is shown superimposed upon the

unbound structure of $G\beta_1\gamma_1$ (1TBG) (Gaudet, Bohm et al. 1996; Sondek, Bohm et al. 1996). Phosducin is in green, $G\beta_1$ in teal, and $G\gamma_1$ from 2TRC is in purple; $G\beta_1\gamma_1$ from 1TBG is in grey. Again, $G\beta\gamma$ does not undergo global conformational changes when bound to its regulator in the visual transduction pathway; the RMSD between the two $G\beta$ molecules is 0.91 Å for $C\alpha$ of residues $G\beta$ 40-340. This structure is more typical of $G\beta\gamma$ signaling complexes, where a $G\beta\gamma$ target interacts both with the “top” face of $G\beta$, and also with the “side” of $G\beta$ through blade interactions.

In summary, much has been learned from structures of G protein signaling components, whether they be the core of the signaling pathways ($G\alpha$, $G\beta$, and $G\gamma$ subunits) or complexes of signaling molecules with their signaling targets.

DESCRIPTION OF DISSERTATION RESEARCH

The dissertation research described in this thesis can be separated into two distinct projects: the structural determination of a complex between a peptide and the $G\beta_1\gamma_2$ heterodimer and preliminary attempts to solve the structure of $G\alpha_{i1}$ bound to fluorescent nucleotides. The first project involves work described earlier in this chapter by Scott et al. concerning the SIRK peptide. Using a doping mutagenesis and rescreening strategy, a peptide similar to the SIRK peptide was derived that had higher affinity for $G\beta_1\gamma_2$. The sequence of this peptide is SIGKAFKILGYPDYD (SIGK). *In vitro* studies with the SIGK peptide indicate that it too can displace $G\alpha_{i1}$ from a heterotrimeric complex and also effectively prevents heterotrimer formation (Ghosh, Peterson et al. 2003). The mechanism by which SIRK/SIGK mediates the dissociation of $G\alpha_{i1}\cdot\text{GDP}$ from $G\beta_1\gamma_2$ is not understood but was suggested to require a conformational change in $G\beta_1\gamma_2$ subunits to account for the enhanced $G\alpha_{i1}$ subunit

dissociation rate in the presence of peptide (Ghosh, Peterson et al. 2003). The structure of the SIGK•G $\beta_{1\gamma 2}$ complex might be used to explain the biological effects seen for the peptide, both in terms of interfering with biological activity assays, and also to explain the dissociation activity of the peptide. In addition, since the peptide was obtained from a phage display, and competed with several other peptides from the display that were unrelated in sequence, it was hoped that the peptide would sample a “hot spot” on G β for protein binding. This analysis of G β as a hot spot for protein interaction might provide insight into the mechanism of G β target binding removed from the context of specific protein•protein interactions.

The second project concerns transitions from inactive to active conformations in the G α subunit. A full introduction to the fluorophore nucleotides used in the present study is included in Chapter V; however, the purpose of the project was to identify new conformations for the G α subunit that might be intermediate between fully activated and fully inactive structures. There is some evidence that the binding of fluorophore nucleotides does not induce the expected conformational changes in G α subunits; a crystallographic snapshot of the conformation found in these nucleotide bound forms might, therefore, allow a glimpse at less-populated conformational states that G α might sample during signaling processes. In addition, the fluorophore nucleotides have altered kinetic properties relative to their unlabeled counterparts; structures of G α bound to these nucleotides might provide structural insights that

could lead to new design of fluorophore nucleotides that have desirable kinetic properties.

Chapter II chronicles the production of $G\beta_1\gamma_2$ heterodimer, formation of the $SIGK \cdot G\beta_1\gamma_2$ complex, and crystallization of that complex.

Chapter III explains the methods used for the structure determination of the $SIGK \cdot G\beta_1\gamma_2$ complex and addresses an interesting result concerning molecular replacement in the structure solution of this complex.

Chapter IV contains a detailed structural analysis of the $G\beta_1\gamma_2 \cdot SIGK$ complex and addresses larger questions about the importance of the “top” face of $G\beta$ as a multiprotein binding site.

Chapter V consists of the methods and results of the first attempt at crystallizing untagged $G\alpha_{i1}$ in the presence of fluorophore nucleotides, and the lessons learned from the experience of crystallizing a copurifying contaminant.

Chapter VI is a summary of the characterization and crystallization of a new construct of $G\alpha_{i1}$, and progress in crystallizing this protein with fluorophore nucleotides.

Chapter VII addresses the conclusions and further directions from the two projects described in Chapters II-VI.

BIBLIOGRAPHY

- Agler, H. L., J. Evans, et al. (2003). "Custom distinctions in the interaction of G-protein beta subunits with N-type (CaV2.2) versus P/Q-type (CaV2.1) calcium channels." J Gen Physiol **121**(6): 495-510.
- Albert, P. R. and L. Robillard (2002). "G protein specificity: traffic direction required." Cell Signal **14**(5): 407-18.
- Berman, D. M., T. Kozasa, et al. (1996). "The GTPase-activating protein RGS4 stabilizes the transition state for nucleotide hydrolysis." J Biol Chem **271**(44): 27209-12.
- Birnbaumer, L. (1992). "Receptor-to-effector signaling through G proteins: roles for beta gamma dimers as well as alpha subunits." Cell **71**(7): 1069-72.
- Bluml, K., W. Schnepf, et al. (1997). "A small region in phosducin inhibits G-protein betagamma-subunit function." Embo J **16**(16): 4908-15.
- Bohm, A., R. Gaudet, et al. (1997). "Structural aspects of heterotrimeric G-protein signaling." Curr Opin Biotechnol **8**(4): 480-7.
- Cabrera-Vera, T. M., J. Vanhauwe, et al. (2003). "Insights into G protein structure, function, and regulation." Endocr Rev **24**(6): 765-81.
- Cismowski, M. J., A. Takesono, et al. (2001). "Receptor-independent activators of heterotrimeric G-proteins." Life Sci **68**(19-20): 2301-8.
- Clackson, T. and J. A. Wells (1994). "In vitro selection from protein and peptide libraries." Trends Biotechnol **12**(5): 173-84.
- Clapham, D. E. and E. J. Neer (1997). "G protein beta gamma subunits." Annu Rev Pharmacol Toxicol **37**: 167-203.
- Coleman, D. E., A. M. Berghuis, et al. (1994). "Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis." Science **265**(5177): 1405-12.
- Coleman, D. E. and S. R. Sprang (1998). "Crystal structures of the G protein Gi alpha 1 complexed with GDP and Mg²⁺: a crystallographic titration experiment." Biochemistry **37**(41): 14376-85.

- Coleman, D. E. and S. R. Sprang (1999). "Structure of G α 1.GppNHp, autoinhibition in a G α protein-substrate complex." J Biol Chem **274**(24): 16669-72.
- DeLano, W. L. (2002). "Unraveling hot spots in binding interfaces: progress and challenges." Curr Opin Struct Biol **12**(1): 14-20.
- DeLano, W. L., M. H. Ultsch, et al. (2000). "Convergent Solutions to Binding at a Protein-Protein Interface." Science **287**(5456): 1279-1283.
- Doering, C. J., A. E. Kisilevsky, et al. (2004). "A single G β subunit locus controls cross-talk between protein kinase C and G protein regulation of N-type calcium channels." J Biol Chem **279**(28): 29709-17.
- Ford, C. E., N. P. Skiba, et al. (1998). "Molecular basis for interactions of G protein betagamma subunits with effectors." Science **280**(5367): 1271-4.
- Gaudet, R., A. Bohm, et al. (1996). "Crystal structure at 2.4 angstroms resolution of the complex of transducin betagamma and its regulator, phosducin." Cell **87**(3): 577-88.
- Gettemans, J., K. Meerschaert, et al. (2003). "A kelch beta propeller featuring as a G beta structural mimic: reinventing the wheel?" Sci STKE **2003**(191): PE27.
- Ghosh, M., Y. K. Peterson, et al. (2003). "Receptor- and nucleotide exchange-independent mechanisms for promoting G protein subunit dissociation." J Biol Chem **278**(37): 34747-50.
- Goubaeva, F., M. Ghosh, et al. (2003). "Stimulation of cellular signaling and G protein subunit dissociation by G protein betagamma subunit-binding peptides." J Biol Chem **278**(22): 19634-41.
- Hamm, H. E. and A. Gilchrist (1996). "Heterotrimeric G proteins." Curr Opin Cell Biol **8**(2): 189-96.
- Harris, B. Z., B. J. Hillier, et al. (2001). "Energetic determinants of internal motif recognition by PDZ domains." Biochemistry **40**(20): 5921-30.
- He, W., L. Lu, et al. (2000). "Modules in the photoreceptor RGS9-1.G β 5L GTPase-accelerating protein complex control effector coupling, GTPase acceleration, protein folding, and stability." J Biol Chem **275**(47): 37093-100.
- Hoey, T., R. O. Weinzierl, et al. (1993). "Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators." Cell **72**(2): 247-60.

- Hou, Y., I. Azpiazu, et al. (2000). "Selective role of G protein gamma subunits in receptor interaction." J Biol Chem **275**(50): 38961-4.
- Hou, Y., V. Chang, et al. (2001). "G Protein beta subunit types differentially interact with a muscarinic receptor but not adenylyl cyclase type II or phospholipase C-beta 2/3." J Biol Chem **276**(23): 19982-8.
- Hynes, T. R., L. Tang, et al. (2004). "Visualization of G protein betagamma dimers using bimolecular fluorescence complementation demonstrates roles for both beta and gamma in subcellular targeting." J Biol Chem **279**(29): 30279-86.
- Inglese, J., L. M. Luttrell, et al. (1994). "Functionally active targeting domain of the beta-adrenergic receptor kinase: an inhibitor of G beta gamma-mediated stimulation of type II adenylyl cyclase." Proc Natl Acad Sci U S A **91**(9): 3637-41.
- Iniguez-Lluhi, J. A., M. I. Simon, et al. (1992). "G protein beta gamma subunits synthesized in Sf9 cells. Functional characterization and the significance of prenylation of gamma." J Biol Chem **267**(32): 23409-17.
- Keleher, C. A., M. J. Redd, et al. (1992). "Ssn6-Tup1 is a general repressor of transcription in yeast." Cell **68**(4): 709-19.
- Koch, W. J., J. Inglese, et al. (1993). "The binding site for the beta gamma subunits of heterotrimeric G proteins on the beta-adrenergic receptor kinase." J Biol Chem **268**(11): 8256-60.
- Krapivinsky, G., M. E. Kennedy, et al. (1998). "Gbeta binding to GIRK4 subunit is critical for G protein-gated K⁺ channel activation." J Biol Chem **273**(27): 16946-52.
- Lambright, D. G., J. P. Noel, et al. (1994). "Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein." Nature **369**(6482): 621-8.
- Lambright, D. G., J. Sondek, et al. (1996). "The 2.0 Å crystal structure of a heterotrimeric G protein." Nature **379**(6563): 311-9.
- Levitzki, A. and S. Klein (2002). "G-protein subunit dissociation is not an integral part of G-protein action." Chembiochem **3**(9): 815-8.
- Li, Y., P. M. Sternweis, et al. (1998). "Sites for Galpha binding on the G protein beta subunit overlap with sites for regulation of phospholipase Cbeta and adenylyl cyclase." J Biol Chem **273**(26): 16265-72.

- Lodowski, D. T., J. A. Pitcher, et al. (2003). "Keeping G proteins at bay: a complex between G protein-coupled receptor kinase 2 and Gbetagamma." Science **300**(5623): 1256-62.
- Loew, A., Y. K. Ho, et al. (1998). "Phosducin induces a structural change in transducin beta gamma." Structure **6**(8): 1007-19.
- Ma, B., H. J. Wolfson, et al. (2001). "Protein functional epitopes: hot spots, dynamics and combinatorial libraries." Curr Opin Struct Biol **11**(3): 364-9.
- Marinissen, M. J. and J. S. Gutkind (2001). "G-protein-coupled receptors and signaling networks: emerging paradigms." Trends Pharmacol Sci **22**(7): 368-76.
- Mirshahi, T., L. Robillard, et al. (2002). "Gbeta residues that do not interact with Galpha underlie agonist-independent activity of K⁺ channels." J Biol Chem **277**(9): 7348-55.
- Mixon, M. B., E. Lee, et al. (1995). "Tertiary and quaternary structural changes in Gi alpha 1 induced by GTP hydrolysis." Science **270**(5238): 954-60.
- Neer, E. J., C. J. Schmidt, et al. (1994). "The ancient regulatory-protein family of WD-repeat proteins." Nature **371**(6495): 297-300.
- Neves, S. R., P. T. Ram, et al. (2002). "G protein pathways." Science **296**(5573): 1636-9.
- Offermanns, S. (2003). "G-proteins as transducers in transmembrane signalling." Prog Biophys Mol Biol **83**(2): 101-30.
- Page, A. M. and P. Hieter (1999). "The anaphase-promoting complex: new subunits and regulators." Annu Rev Biochem **68**: 583-609.
- Pryer, N. K., N. R. Salama, et al. (1993). "Cytosolic Sec13p complex is required for vesicle formation from the endoplasmic reticulum in vitro." J Cell Biol **120**(4): 865-75.
- Richardson, M. and J. D. Robishaw (1999). "The alpha2A-adrenergic receptor discriminates between Gi heterotrimers of different betagamma subunit composition in Sf9 insect cell membranes." J Biol Chem **274**(19): 13525-33.
- Roberts, D. J. and M. Waelbroeck (2004). "G protein activation by G protein coupled receptors: ternary complex formation or catalyzed reaction?" Biochem Pharmacol **68**(5): 799-806.

- Robishaw, J. D. and C. H. Berlot (2004). "Translating G protein subunit diversity into functional specificity." Curr Opin Cell Biol **16**(2): 206-9.
- Rondard, P., T. Iiri, et al. (2001). "Mutant G protein alpha subunit activated by Gbeta gamma: a model for receptor activation?" Proc Natl Acad Sci U S A **98**(11): 6150-5.
- Ross, E. M. and T. M. Wilkie (2000). "GTPase-activating proteins for heterotrimeric G proteins: regulators of G protein signaling (RGS) and RGS-like proteins." Annu Rev Biochem **69**: 795-827.
- Sankaran, B., J. Osterhout, et al. (1998). "Identification of a structural element in phospholipase C beta2 that interacts with G protein betagamma subunits." J Biol Chem **273**(12): 7148-54.
- Schwindinger, W. F., K. S. Betz, et al. (2003). "Loss of G protein gamma 7 alters behavior and reduces striatal alpha(olf) level and cAMP production." J Biol Chem **278**(8): 6575-9.
- Schwindinger, W. F. and J. D. Robishaw (2001). "Heterotrimeric G-protein betagamma-dimers in growth and differentiation." Oncogene **20**(13): 1653-60.
- Scott, J. K., S. F. Huang, et al. (2001). "Evidence that a protein-protein interaction 'hot spot' on heterotrimeric G protein betagamma subunits is used for recognition of a subclass of effectors." Embo J **20**(4): 767-76.
- Simon, M. I., M. P. Strathmann, et al. (1991). "Diversity of G proteins in signal transduction." Science **252**(5007): 802-8.
- Simonds, W. F. and J. H. Zhang (2000). "New dimensions in G protein signalling: G beta 5 and the RGS proteins." Pharm Acta Helv **74**(2-3): 333-6.
- Skelton, N. J., Y. M. Chen, et al. (2001). "Structure-function analysis of a phage display-derived peptide that binds to insulin-like growth factor binding protein 1." Biochemistry **40**(29): 8487-98.
- Slep, K. C., M. A. Kercher, et al. (2001). "Structural determinants for regulation of phosphodiesterase by a G protein at 2.0 Å." Nature **409**(6823): 1071-7.
- Snow, B. E., A. M. Krumins, et al. (1998). "A G protein gamma subunit-like domain shared between RGS11 and other RGS proteins specifies binding to Gbeta5 subunits." Proc Natl Acad Sci U S A **95**(22): 13307-12.
- Sondek, J., A. Bohm, et al. (1996). "Crystal structure of a G-protein beta gamma dimer at 2.1 Å resolution." Nature **379**(6563): 369-74.

- Sprang, S. R. (1997). "G protein mechanisms: insights from structural analysis." Annu Rev Biochem **66**: 639-78.
- Tang, W. J. and A. G. Gilman (1991). "Type-specific regulation of adenylyl cyclase by G protein beta gamma subunits." Science **254**(5037): 1500-3.
- Tesmer, J. J., D. M. Berman, et al. (1997). "Structure of RGS4 bound to AlF₄--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis." Cell **89**(2): 251-61.
- Tesmer, J. J., R. K. Sunahara, et al. (1997). "Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS." Science **278**(5345): 1907-16.
- Ueda, N., J. A. Iniguez-Lluhi, et al. (1994). "G protein beta gamma subunits. Simplified purification and properties of novel isoforms." J Biol Chem **269**(6): 4388-95.
- Violot, S., S. S. Hong, et al. (2003). "The human polycomb group EED protein interacts with the integrase of human immunodeficiency virus type 1." J Virol **77**(23): 12507-22.
- Wall, M. A., D. E. Coleman, et al. (1995). "The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2." Cell **83**(6): 1047-58.
- Wall, M. A., B. A. Posner, et al. (1998). "Structural basis of activity and subunit recognition in G protein heterotrimers." Structure **6**(9): 1169-83.
- Weng, G., J. Li, et al. (1996). "Gbeta subunit interacts with a peptide encoding region 956-982 of adenylyl cyclase 2. Cross-linking of the peptide to free Gbetagamma but not the heterotrimer." J Biol Chem **271**(43): 26445-8.
- Wettschureck, N., A. Moers, et al. (2004). "Mouse models to study G-protein-mediated signaling." Pharmacol Ther **101**(1): 75-89.
- Yan, K. and N. Gautam (1996). "A domain on the G protein beta subunit interacts with both adenylyl cyclase 2 and the muscarinic atrial potassium channel." J Biol Chem **271**(30): 17597-600.

CHAPTER TWO

Production of $G\beta_1\gamma_2$ and the $G\beta_1\gamma_2$ •SIGK Complex, and Crystallization of the $G\beta_1\gamma_2$ •SIGK Complex

The structure of the $G\beta_1\gamma_2$ •SIGK complex was undertaken in the context of three papers concerning the biological activity of the SIGK peptide as described in Chapter One. In 2001 Scott et al. described an experiment in which phage displays of randomized peptide libraries were screened for binding to $G\beta_1\gamma_2$ dimers (Scott, Huang et al. 2001). The peptides that bound $G\beta_1\gamma_2$ and were pulled out of their assay could be loosely classified into four unrelated groups based on amino acid sequence (Figure 2-1). One of these groups included a linear peptide (the “SIRK” peptide) with the sequence SIRKALNILGYPDYD. Peptides belonging to all four groups competed with each other with a range of affinities for binding to $G\beta_1\gamma_2$, suggesting that all of the clones isolated from the phage display screen shared a common binding site on $G\beta_1\gamma_2$ (Scott, Huang et al. 2001). In addition, SIRK inhibited the activation of PLC β_2 by $G\beta\gamma$ subunits, and also inhibited the activation of PI3K by $G\beta\gamma$ (Scott, Huang et al. 2001). However, SIRK peptide was a poor inhibitor of adenylyl cyclase type I and voltage-gated Ca^{++} channel pathways, indicating it was not a global inhibitor of $G\beta\gamma$ mediated signaling pathways (Scott, Huang et al. 2001). A subsequent study in which SIRK was modified with either myristate or a cell permeation sequence showed that SIRK was able to activate ERK 1 and 2 isoforms in RASM cells (Goubaeva, Ghosh et al. 2003). Presumably this effect is due to the

ability of SIRK peptide to break apart $G\alpha\beta\gamma$ heterotrimers, thereby freeing $G\beta\gamma$ to signal through downstream effectors.

Using a doping mutagenesis and rescreening strategy, a peptide similar to the SIRK peptide was derived that had higher affinity for $G\beta_1\gamma_2$. The sequence of this peptide is SIGKAFKILGYPDYD (SIGK). *In vitro* studies with the SIGK peptide indicate that it too can displace $G\alpha_{i1}$ from a heterotrimeric complex and also effectively prevents heterotrimer formation (Ghosh, Peterson et al. 2003). The mechanism by which SIRK/SIGK mediates the dissociation of $G\alpha_{i1}\cdot\text{GDP}$ from $G\beta_1\gamma_2$ is not understood but was suggested to require a conformational change in $G\beta_1\gamma_2$ subunits to account for the enhanced $G\alpha_{i1}$ subunit dissociation rate in the presence of peptide (Ghosh, Peterson et al. 2003).

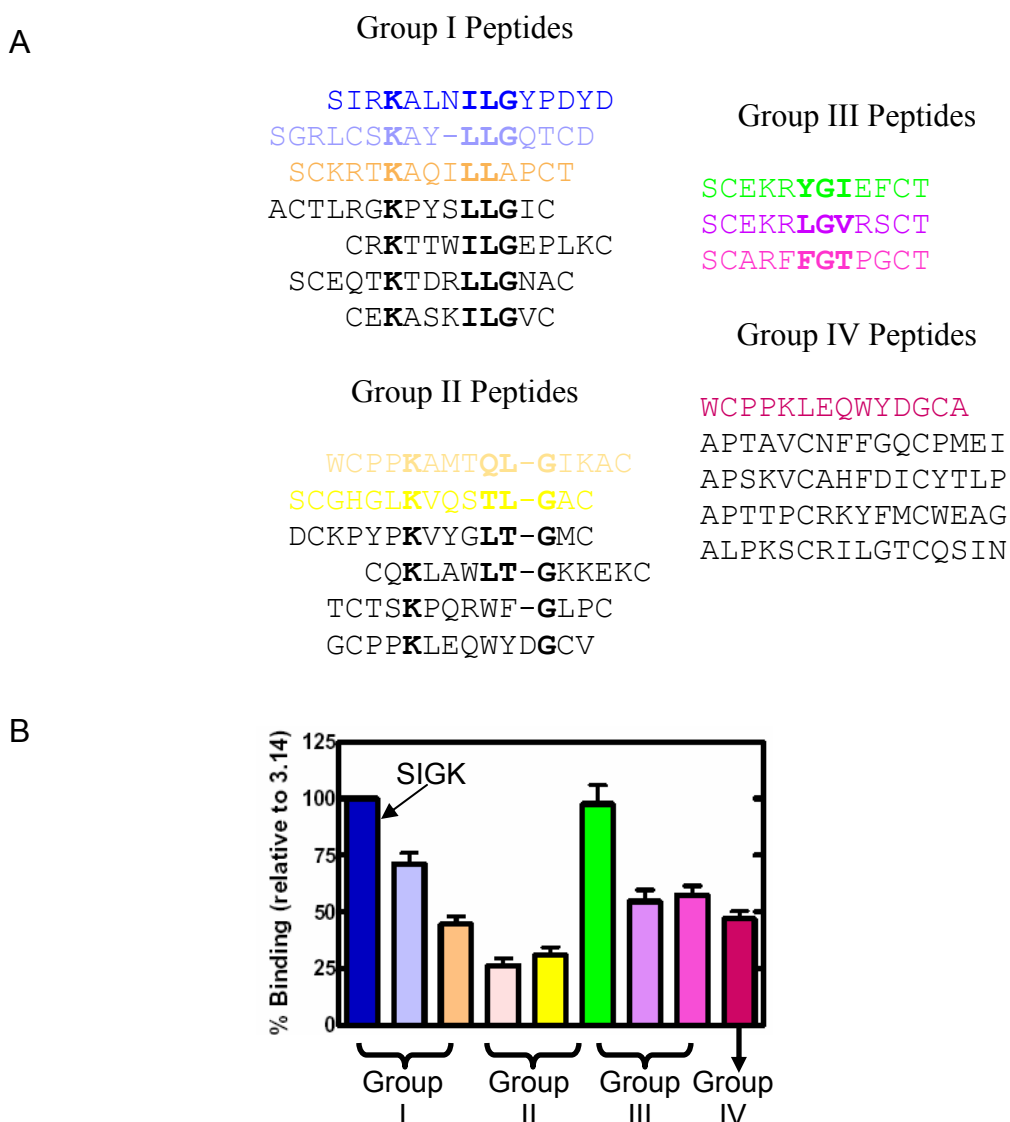


Figure 2-1. Peptides from phage display. (A) is adapted from (Scott, Huang et al. 2001). (B) is from unpublished data from T. Bonacci and A.V. Smrcka. (A) Representative peptides for groups I-IV. Conserved residues are shown in **bold**. For all peptides besides SIRK, there are N terminal and C terminal cysteines to induce disulfide constraints. These cysteines are not shown in the group IV sequences. (B) Using a phage display binding assay (described in Chapter Four), peptides were screened for their ability to bind to immobilized $G\beta\gamma$ subunits. When data is available for the peptides in (A), their sequence is colored as in (B). All binding is shown as a relative percentage of SIGK binding to $G\beta\gamma$.

The goal of the current project was to obtain diffraction quality crystals of the SIGK•G $\beta_1\gamma_2$ complex and solve the crystal structure. The structure obtained might allow both an explanation of the biological effects of the SIGK peptide and provide insight into the molecular characteristics of the “hot spot” sampled on G β by the peptides obtained from the phage display screen. Chapter Two is focused on the production of G $\beta_1\gamma_2$ and SIGK peptide and crystallization of the resultant complex.

MATERIALS AND METHODS

Production and Amplification of G β_1 and G γ_2 Baculoviral Constructs

Baculoviruses harboring cDNA for wild-type bovine G β_1 , N terminally (His)₆-tagged bovine G γ_2 , and N terminally (His)₆-tagged bovine G γ with the C68S mutation were obtained from Alfred Gilman (UTSWMC) (Iniguez-Lluhi, Simon et al. 1992).

Sequences of the encoded genes are shown in Figure 2-2. Generally, the protocols for expression of G $\beta\gamma$ using these viruses are well documented (Iniguez-Lluhi, Simon et al. 1992; Ueda, Iniguez-Lluhi et al. 1994; Kozasa and Gilman 1995; Wall, Coleman et al. 1995); however, a summary of the procedures used for this project are outlined below. Sf9 cells (*Spodoptera frugiperda*, Invitrogen) cultured in IPL-41 media without serum (Invitrogen-Gibco-BRL) but supplemented with 1% Pluronic F68 (Invitrogen-Gibco-BRL) and 1% lipid concentrate (Invitrogen-Gibco-BRL) were used to propagate virus using instruction manuals from Invitrogen. High 5 cells (BRI-TN-5B1-4 cell line, derived from *Trichopulsia ni*, also from Invitrogen-Gibco-

BRL) grown in Insect Express media without supplements (Fisher Scientific) were used for testing baculovirus batches for protein expression, and also used for large-scale protein expression.

> Bovine G β_1

MSELDQLRQEAEQLKNQIRDARKACADATLSQITNNIDPVGRIQMRTTRTLRGHLAKIYAMHWGTDSRLLVSASQ
DGKLIWDSYTTNKVHAIPLRSSWVMTCAAPSGNYVACGGLDNICSIYNLKTREGNVRVSRELAGHTGYLSCCRF
LDDNQIVTSSGDTTCALWDIETGQQTTFGTGTDVMSLSLAPDTRLFVSGACDASAKLWDVREGMCRQTFTGH
ESDINAICFFPNGNAFATGSDDATCRLFDLRADQELMTYSHDNIICGITSVSFSKSGRLLLAGYDDFNCNVWDALK
ADRAGVLAGHDNRVSLGVTDDGMAVATGSWDSFLKIWN

Residues: 340

Molecular Weight: 37.3 kDa

Predicted pI: 5.6

>Bovine G γ_2

MASNNTASIAQARKLVEQLKMEANIDRIKVSAAAADLMAYCEAHAKEDPLLTPVPAENPFREKKFF(C/S)AIL

Residues: 71

Molecular Weight: 7.8 kDa

Predicted pI: 7.8

SIGK peptide

SIGKAFKILGYPDYD

Residues: 15

Molecular Weight: 1.7 kDa

Predicted pI: 5.9

Figure 2-2. Sequences of proteins used in the current project. For G γ_2 , the position of the C68S mutation is shown. Documentation of the source DNA for G β_1 and both G γ viruses can be found in (Iniguez-Lluhi, Simon et al. 1992).

Expression and Purification of the G $\beta_1\gamma_2$ Heterodimer

G $\beta_1\gamma_2$ was purified with a modified method based on that of Kosaza and Gilman (Kozasa and Gilman 1995). All steps were carried out at 4 °C. Cells were harvested 60 hours post-infection by centrifugation at 2600g, then resuspended in 50 ml of lysis buffer (20 mM HEPES, pH 8, 150 mM NaCl, 5 mM β -ME, 1mM EDTA, 1 mL Sigma protease inhibitor cocktail P-2714) per liter of cell culture. Cells were lysed by sonication (5 sec on, 5 sec off, 5 minutes total sonication time) and centrifuged at

2600g to pellet the membranes. Resuspension and homogenization of membranes was accomplished by douncing in 100 mL lysis buffer. The membranes were solubilized by adding 1% Lubrol ($C_{12}E_{10}$, Sigma) with stirring and the resultant solution clarified by ultracentrifugation at 125,000g. The supernatant was loaded onto Ni-NTA agarose (Qiagen) equilibrated with lysis buffer+1% Lubrol. The column was washed and the Lubrol exchanged for sodium cholate using buffers Ni-A (20 mM HEPES, pH 8, 0.4 M NaCl, 5 mM β -ME, 0.5% Lubrol, 0.15% cholate) and Ni-B (20 mM HEPES pH 8, 0.1M NaCl, 5 mM β -ME, 0.25% Lubrol, 0.3% cholate), respectively. $G\beta_{1\gamma_2}$ was eluted in Ni-C (20 mM HEPES pH 8, 0.01 M NaCl, 5 mM β -ME, 1% cholate, 200 mM imidazole). A final detergent exchange into CHAPS was accomplished during anion exchange chromatography. The eluate was loaded onto a HiTrap Q (Amersham Biosciences) column pre-equilibrated with QA (20 mM HEPES, pH 8, 5 mM β -ME, 0.7% CHAPS, 1 mM EDTA). $G\beta_{1\gamma_2}$ was eluted in a linear gradient using QB (QA+1M NaCl). Fractions containing $G\beta_{1\gamma_2}$ were analyzed by SDS-PAGE and pooled. Gel filtration was performed using a tandem Sephadex 75: Sephadex 200 column (Amersham Biosciences) equilibrated with buffer GF+CHAPS (20 mM HEPES, pH 8, 150 mM NaCl, 10 mM β -ME, 1 mM EDTA, 0.7% CHAPS). When β -octylglucoside was used for crystallization trials, the gel filtration step was run using GF+ β OG buffer (20 mM HEPES, pH8, 150 mM NaCl, 10 mM β -ME, 1 mM EDTA, 1% β -OG). Identical methods were used to purify $G\beta_{1\gamma_2}$ heterodimers containing the C68S mutation, except that all detergents were

eliminated from the buffers and the supernatant from the sonication step was ultracentrifuged and loaded directly onto Ni-NTA agarose.

Production of SIGK Peptide

SIGK peptide (SIGKAFKILGYPDYD) was synthesized by the Protein Chemistry Technology Center at UTSWMC. No modifications were made to the peptide termini; purification was by reverse phase-HPLC chromatography on a Vydac C4 semi-preparative column.

Formation and Crystallization of $G\beta_{1\gamma_2}$ •SIGK Complex

SIGK peptide was added to $G\beta_{1\gamma_2}$ in 1.5 molar excess, and the $G\beta_{1\gamma_2}$ •SIGK complex was used at 7 mg mL⁻¹ for crystallization. Crystals were grown by vapor diffusion using equal volumes (2 μ l) of protein and reservoir solution (15-17% PEG 4000, 100 mM HEPES, pH 7.5, 0.01-0.05 M Na-Acetate, 10% glycerol) at 20 °C. Crystals attained dimensions of 150 μ x 50 μ x 20 μ within one week.

Stabilization and Freezing of Crystals

Crystals were cryoprotected in 15% glycerol and frozen in liquid nitrogen using standard methods. In brief, crystals were harvested from mother liquor and placed into a stabilization solution containing 20% PEG 4000, 100 mM HEPES pH 7.5, 0.05 M Na-Acetate, and 15 % glycerol, and allowed to soak for 1-5 minutes. Cryoloops of different diameters (usually 50-100 microns) (Hampton Research) were used to

harvest, transfer, and mount the crystals. The crystals were taken from the cryoprotectant solution and either flash frozen directly into liquid nitrogen for storage, or placed quickly into a cold nitrogen stream (110 K) for screening. Mounted and cryoprotected crystals appeared clear and glassy, indicating sufficient cryoprotection; in addition, no ice spots were detected for most crystal diffraction patterns.

Crystal Screening and Data Collection using Synchrotron Sources

Native crystals of $G\beta_{1\gamma_2}$ •SIGK were screened at Advanced Light Source (ALS) beamlines 8.2.1 and 8.2.2 (Berkeley, CA) and at the Advanced Photon Source (APS) beamline BM-19 (Chicago, IL). A dataset from ALS 8.2.2 was used to determine the structure. Over 100 crystals were screened; diffraction limits varied from 7Å to the 2.7Å dataset used for structure determination. Diffraction data were indexed, integrated, and scaled using the software package HKL2000 (Minor 1997).

RESULTS AND DISCUSSION

Expression and Purification of $G\beta_{1\gamma_2}$

The expression and purification of $G\beta_{1\gamma_2}$ and $G\beta_{1\gamma_2}$ (C68S) has been described (Iniguez-Lluhi, Simon et al. 1992; Ueda, Iniguez-Lluhi et al. 1994; Kozasa and Gilman 1995; Wall, Coleman et al. 1995). The task at hand was to obtain crystallographic quantities of the wild type $G\beta_{1\gamma_2}$ and $G\beta_{1\gamma_2}$ (C68S) protein in order to

screen for conditions for a SIGK•G $\beta\gamma$ complex. Unpublished information from A. V. Smrcka's lab indicated that the SIGK complex preferentially formed in the presence of the prenylation on the G γ_2 subunit, suggesting that the wild type complex might be the better crystallographic target. However, this data could not be replicated using alternative methods, and conditions for crystallizing G $\beta\gamma$ complex solubilized in detergent were unknown. Therefore, both forms of the G $\beta\gamma$ •SIGK complex were isolated in the hope of quickly obtaining diffraction quality crystals.

On the following six pages are gels and chromatograms showing the relative purity at different stages of the G $\beta\gamma$ (C68S) and wild type G $\beta\gamma$ protein preps (Figures 2-3 through 2-9).

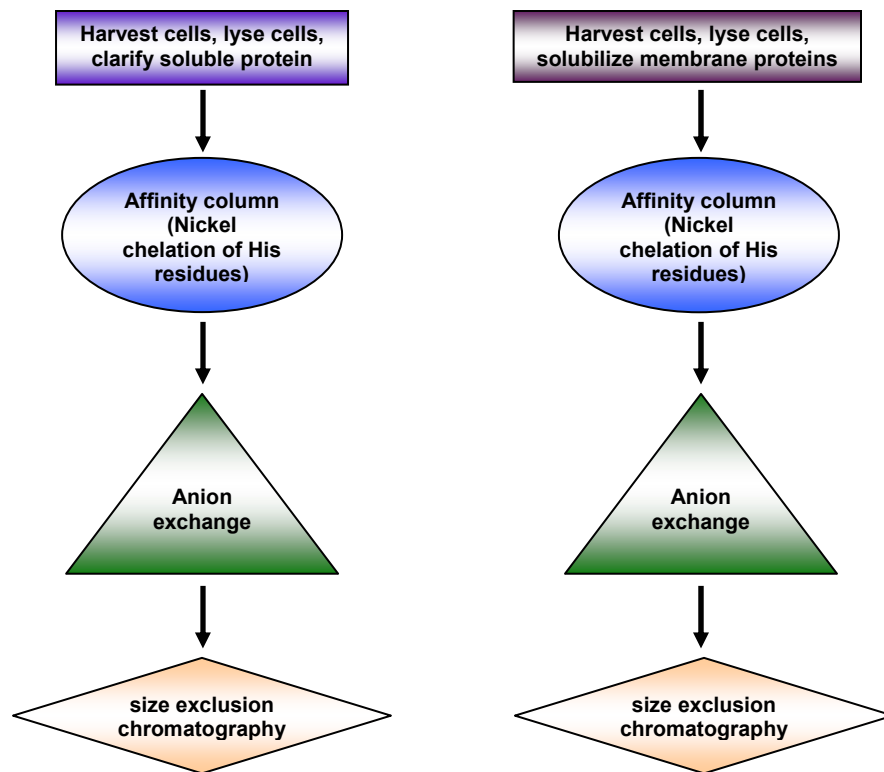
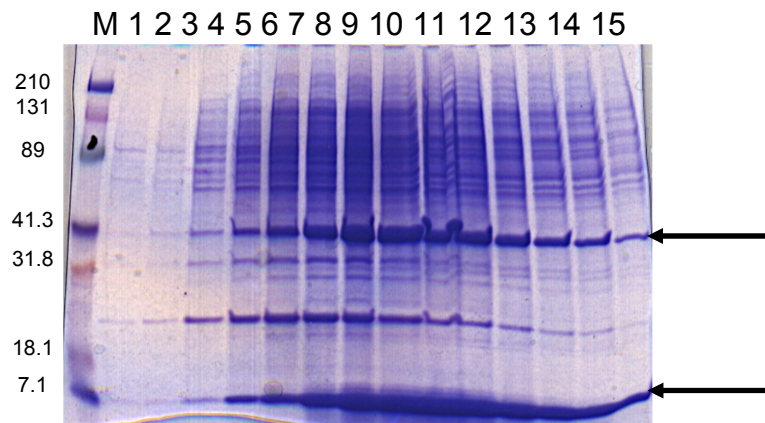


Figure 2-3. Flow chart of $G\beta_1\gamma_2$ (C68S) and $G\beta_1\gamma_2$ (wild-type) purifications. On the left is shown a protocol for $G\beta\gamma$ containing the C68S mutation; on the right is a similar protocol for wild-type $G\beta\gamma$. The only substantive difference is the source of the supernatant (soluble or membrane-associated and solubilized).

Figure 2-4. Fractions from Ni-NTA, $G\beta_1\gamma_2$ (C68S). 5 μ L of 2 mL fractions from an imidazole elution are shown on a Coomassie-stained SDS-PAGE gel. The marker on this gel is the Kaleidoscope prestained standard from Bio-Rad; molecular weights are shown from 7.1 to 210 kDa. All 15 lanes are eluate fractions from the Ni-NTA resin. Arrows point to $G\beta_1$ and $G\gamma_2$ (C68S) proteins. The expected molecular weight for $G\beta_1$ is 37.3 kDa; for $G\gamma_2$ it is 7.8 kDa.



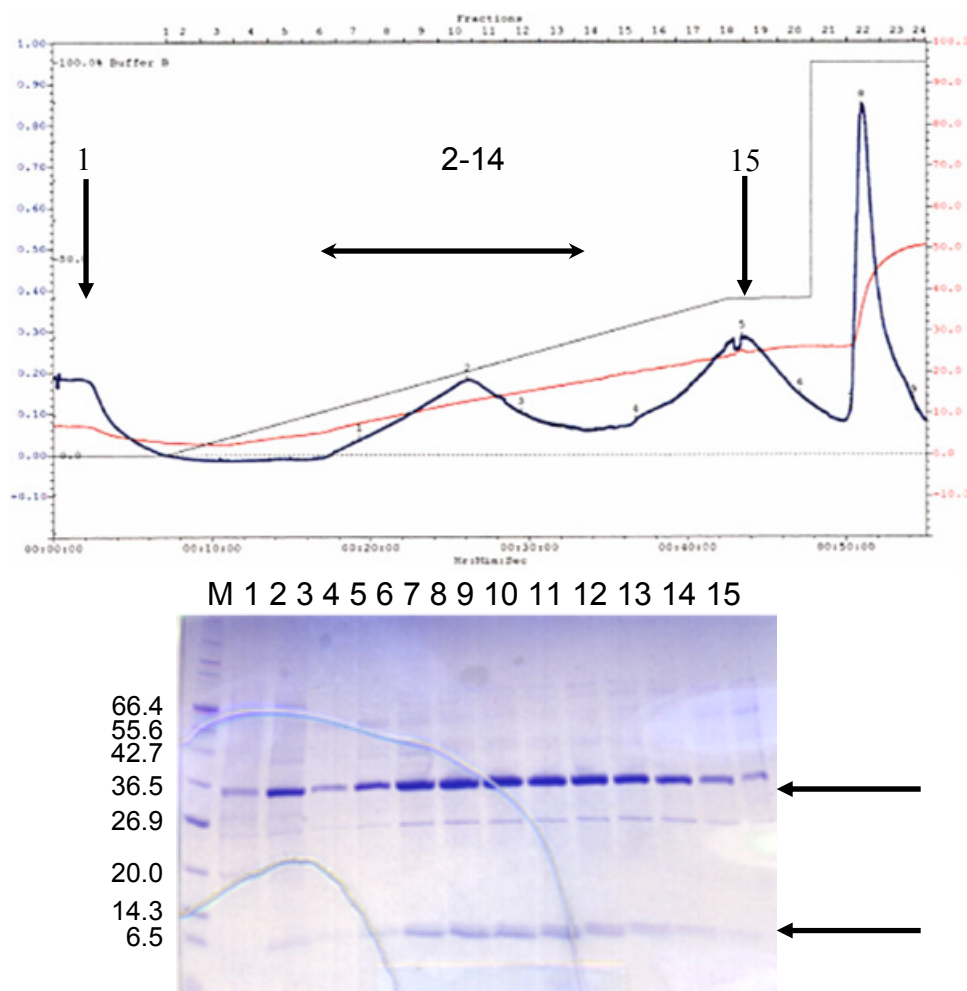


Figure 2-5. Purification from Hitrap Q anion exchange chromatography, $G\beta_1\gamma_2(C68S)$. 10 μ L of 1 mL fractions are shown on a Coomassie-stained SDS-PAGE gel. The marker shown is the Broad Range marker (2.3-212 kDa) from NEB; only the mid-range markers (10-70 kDa) are labeled. The protein samples are as follows: lane 1, flowthrough; lanes 2-14, eluate fractions from 6-13 over the NaCl gradient; lane 15, peak fraction of the second peak (fraction 18). The $G\beta_1\gamma_2(C68S)$ protein elutes from the column in 100-200 mM NaCl, corresponding to the first peak of the chromatogram.

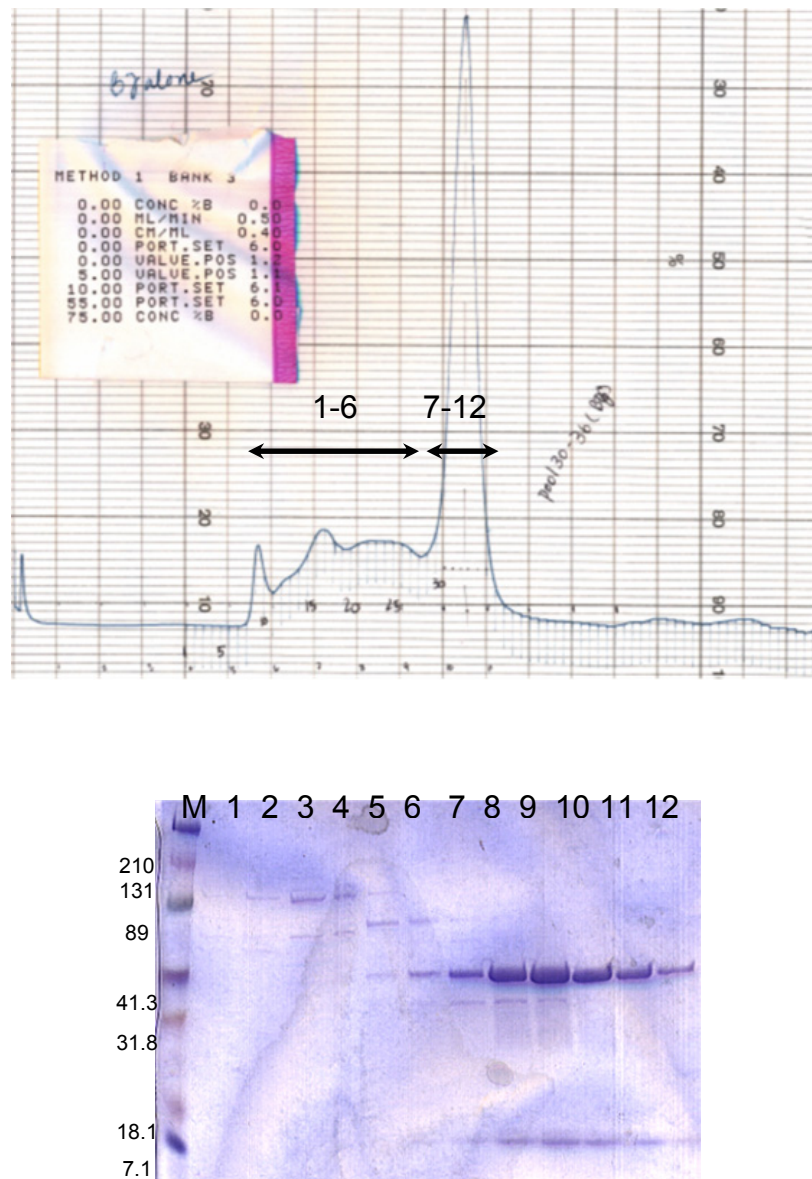


Figure 2-6. Chromatograph and fractions from Sephadex 200:75, $G\beta_{1\gamma_2}(C68S)$. A Coomassie-stained SDS-PAGE gel showing 5 μ L of 0.5 mL fractions is shown. After gel filtration, a slight contaminant at around 35 kDa can be seen in the fractions.

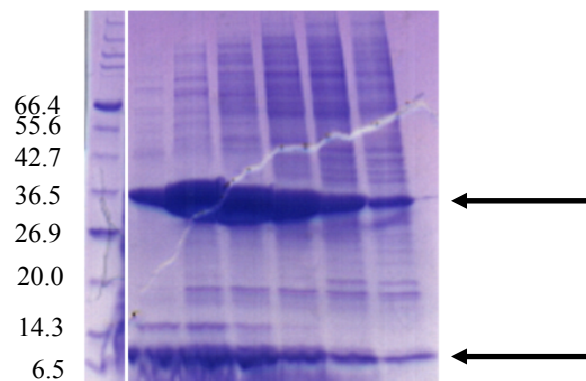
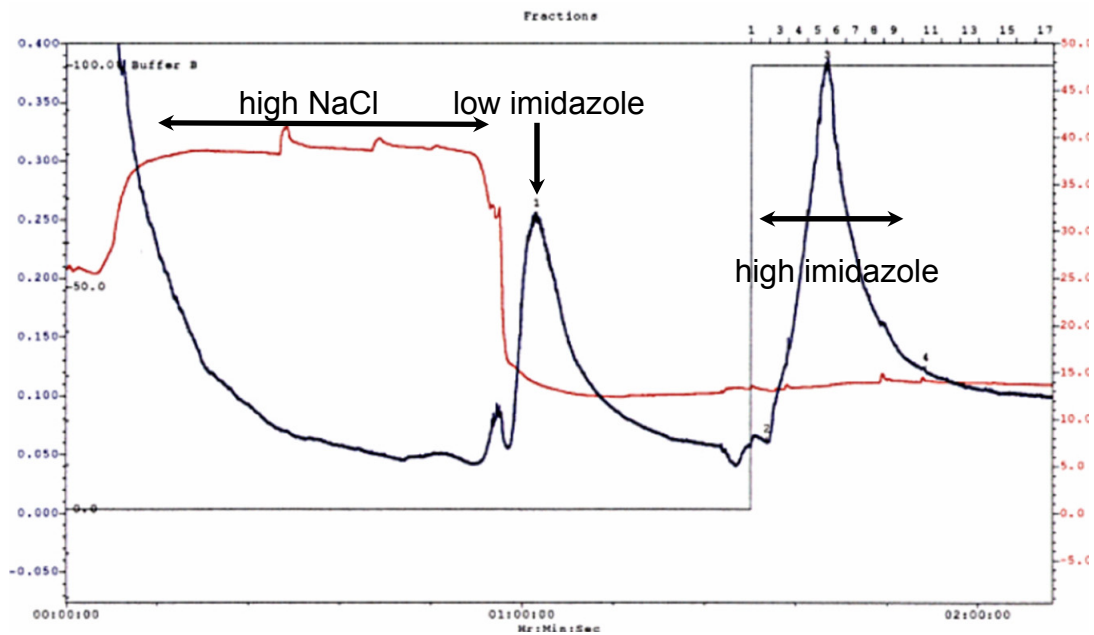


Figure 2-7. Chromatograph and fractions from Ni-NTA, wild type G $\beta_1\gamma_2$. 10 μ L of 5 mL fractions are loaded onto a Coomassie-stained SDS-PAGE gel. Molecular weight marker is from NEB. Although the high salt and low imidazole washes are shown on the chromatogram, only fractions from the high imidazole elution peak are shown on the gel. See Figure 2-8 for an example of the proteins in the high salt and low imidazole washes. The distortion at the bottom of the gel is due to the detergent present in the elution buffer (sodium cholate).

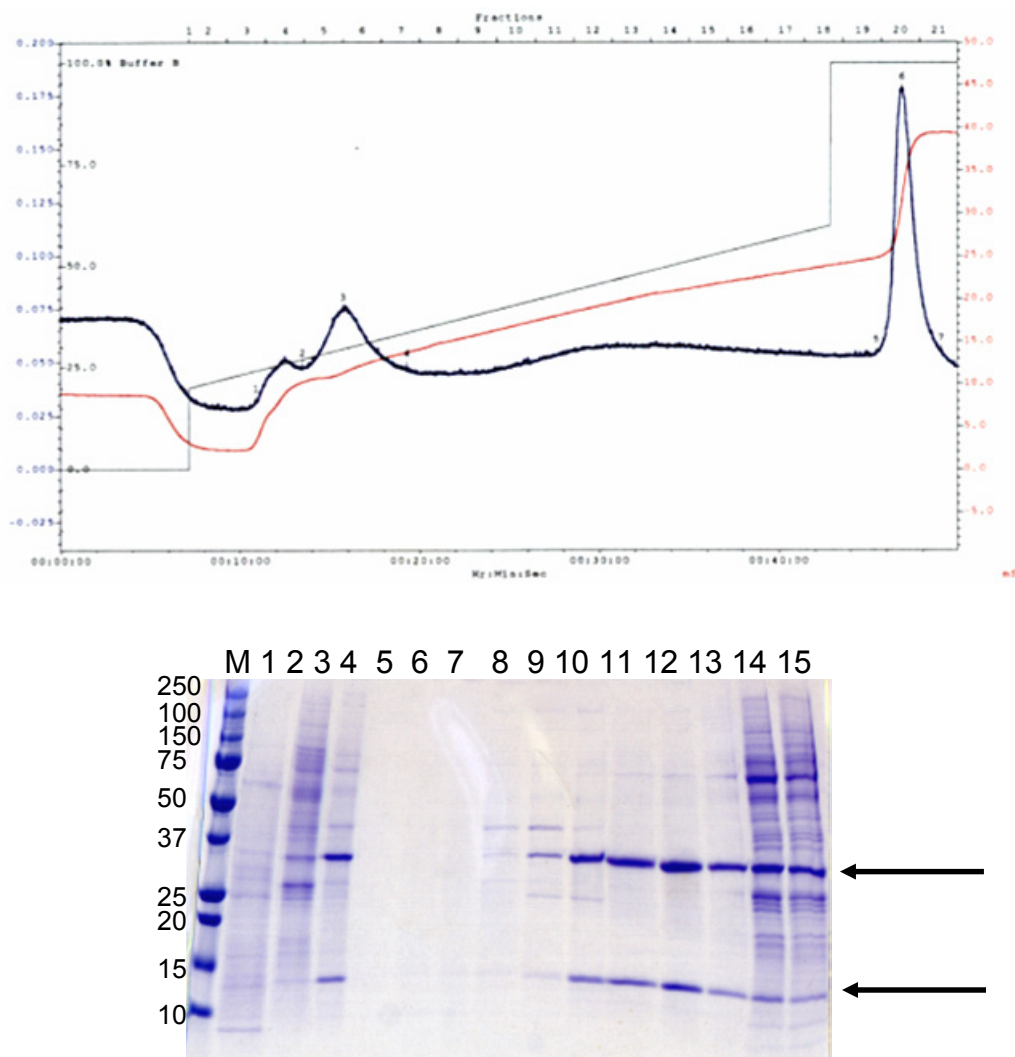


Figure 2-8. Chromatograph and fractions from HiTrapQ, $G\beta_1\gamma_2$. The first four lanes of the gel contain marker (Precision Plus Protein Standards, Bio-Rad, weights shown in kDa), (1) Nickel high salt elute, (2) Nickel low imidazole elute, and (3) Nickel elution. The rest of the gel (4-15) covers the shoulder and first peak shown on the chromatogram, which begins at around 100 mM NaCl. The band at ~13 kDa on the gel could be $G\gamma$, although the molecular weight for this protein is 8 kDa and the discrepancy on this gel is unexplained.

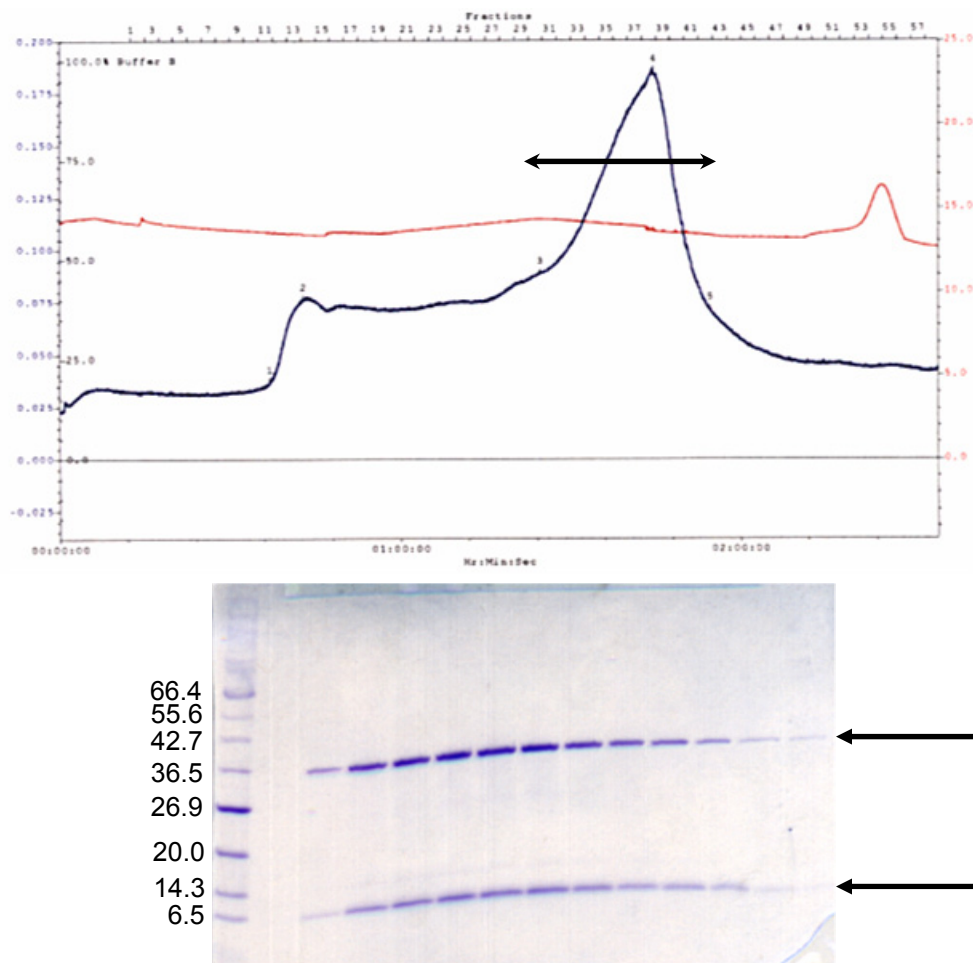


Figure 2-9. Chromatograph and fractions from Sephadex 200:75, G $\beta_1\gamma_2$. The column was run in either CHAPS or β -OG, depending on the crystallographic experiment. This example is from an experiment done in CHAPS buffer, but the results are similar if run in β -OG. 5 μ L of 0.5 mL fractions are shown on a Coomassie-stained SDS-PAGE gel. Although the eluant peak is not wholly symmetric, the fractions do not seem to be contaminated with copurifying proteins.

Typical yields of protein were 1 mg of purified heterodimers per liter of cell culture, both for heterodimers containing wild-type $G\beta_{1\gamma_2}$ and for heterodimers containing the C68S mutation. The purity of the wild-type heterodimers seemed to be higher, but both forms of complex were utilized for initial crystal screens.

Crystallization of the $G\beta_{1\gamma_2}$ •SIGK Complex

Wild-type and C68S mutant heterodimers were both initially used to screen for $G\beta_{1\gamma_2}$ •SIGK complex crystals. The following sparse matrix screens were set up in initial trials: Hampton I and II (Hampton Research) and Wizard I and II (Emerald Biosciences), at 20 °C. Both wild type $G\beta\gamma$ and $G\beta\gamma$ (C68S) mixed with SIGK peptide produced small crystals from conditions in the initial screens (Figures 2-10 through 2-13). The wild-type heterodimer complex formed in the presence of β -OG produced crystals from a screen condition utilizing Peg 4K (Figure 2-11), as did the same complex formed in the presence of CHAPS detergent (Figure 2-12). The C68S heterodimer complex produced crystals from an unrelated screen condition, utilizing K/Na^+ tartrate as the precipitant (Figure 2-10). Attempts were made to optimize all three of these crystallization experiments simultaneously by systematically varying protein concentration, precipitant concentration, pH, and additive concentration when appropriate. Microseeding and macroseeding experiments were also undertaken in an attempt to grow larger crystals as well. However, optimization of the C68S heterodimer complex crystals proceeded slowly, as did optimization of wild type heterodimers complex solubilized in β -OG. The wild-

type heterodimer solubilized in CHAPS produced crystals of the highest quality.

Initial crystals of the $G\beta_1\gamma_2$ •SIGK complex in CHAPS were small and practically two-dimensional; optimization proceeded mainly upon the observation that the addition of glycerol produced fewer and larger crystals. Finally reproducible and diffraction-quality crystals of the $G\beta_1\gamma_2$ •SIGK complex solubilized in CHAPS were obtained, although the crystals were still quite thin ($150\mu \times 50\mu \times 20\mu$) (Figure 2-13). All data sets were collected on wild-type $G\beta_1\gamma_2$ •SIGK complexes solubilized in CHAPS.

Initial crystal screening and attempts at data collection were carried out at 100K on a rotating anode X-ray generator with an R-axis IV image plate detector at UTSWMC. The cryoprotected $G\beta_1\gamma_2$ •SIGK crystals were visible in the cryoloops in a clear glassy state, indicating proper cryoprotection at 15% glycerol. However, diffraction data collected for 60 min exposure time showed no diffraction at all or limited diffraction to 10 Å or less, indicating that these crystals diffracted weakly and would have to be screened using synchrotron radiation. This precluded extensive cryoprotectant screening trials at home. Cryoprotection is essential for diffraction experiments at synchrotron sources, due to the reduction in the rate of radiation decay seen in cryoprotected crystals (Rodgers 1994). In addition, cryoprotection allows for the technique of crystal annealing to be utilized, in which a frozen crystal is allowed to warm and then refrozen. This technique has been shown to improve the mosaicity and even the useful resolution of some protein crystals (Kriminski, Caylor et al. 2002). It is beneficial to attempt cryoprotection in several different types of

chemicals (small molecular weight PEG solutions, sugars, cryosalts (Rubinson, Ladner et al. 2000), or oils (Kwong 1999; Ribolidi-Tunncliffe 1999)) to ascertain the optimal conditions for cryoprotection (Rodgers 1994). Limited trials of MPD, PEG 400, and sucrose at the synchrotron did not yield improvement in diffraction quality or mosaicity measurements on the $G\beta_1\gamma_2$ •SIGK crystals. Data collection protocols and structure solution of the $G\beta_1\gamma_2$ •SIGK complex will be outlined in more detail in Chapter Three.

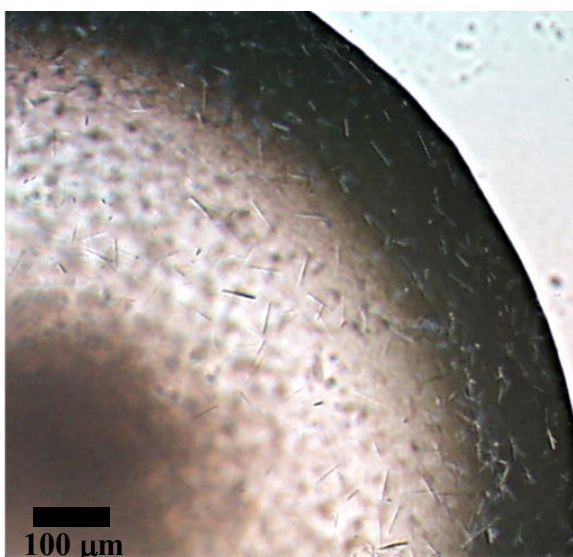


Figure 2-10. Initial crystals of the $G\beta_{1\gamma_2}(C68S)\cdot SIGK$ complex. These crystals were obtained in the condition 0.1M HEPES pH 7.5, 0.8M K/Na Tartrate (Hampton Screen I #29).

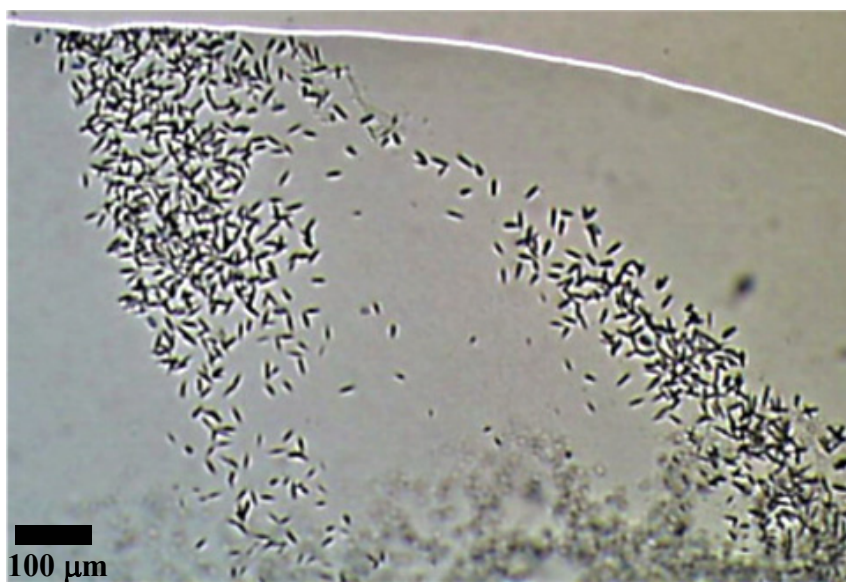


Figure 2-11. Initial crystals of the $G\beta_{1\gamma_2}\cdot SIGK$ complex solubilized in β -OG. These crystals were obtained from a solution containing 20% PEG 4K, 0.1M HEPES pH 7.5, 10% isopropanol (Hampton Screen I condition #41).

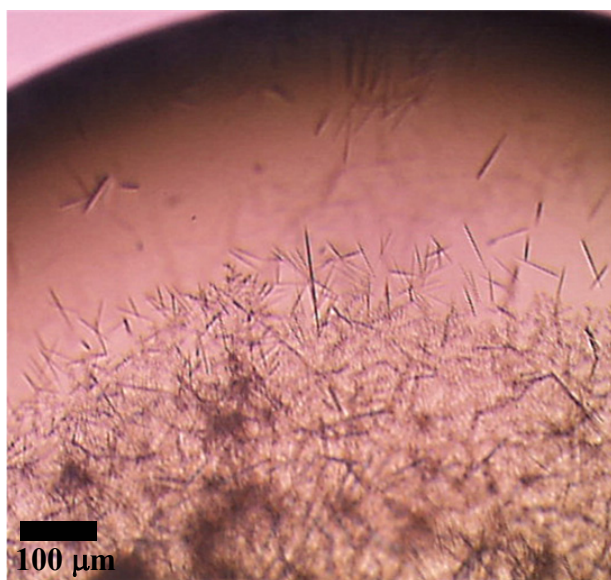


Figure 2-12. Initial crystals of the $G\beta_{1\gamma_2}$ •SIGK complex solubilized in CHAPS. These crystals were obtained from solutions containing 0.1M Tris pH 8.5, 30% PEG 4K, 0.2M Na-Acetate (Hampton Screen I # 22).



Figure 2-13. Optimized crystals of the $G\beta_{1\gamma_2}$ •SIGK complex solubilized in CHAPS. These crystals came from condition ranges: 16.5-16.7% PEG 4K, 0.02-0.1 M Na-Acetate, 10-15% glycerol, and 0.1M Tris pH 8.5.

BIBLIOGRAPHY

- Ghosh, M., Y. K. Peterson, et al. (2003). "Receptor- and nucleotide exchange-independent mechanisms for promoting G protein subunit dissociation." J Biol Chem **278**(37): 34747-50.
- Goubaeva, F., M. Ghosh, et al. (2003). "Stimulation of cellular signaling and G protein subunit dissociation by G protein betagamma subunit-binding peptides." J Biol Chem **278**(22): 19634-41.
- Iniguez-Lluhi, J. A., M. I. Simon, et al. (1992). "G protein beta gamma subunits synthesized in Sf9 cells. Functional characterization and the significance of prenylation of gamma." J Biol Chem **267**(32): 23409-17.
- Kozasa, T. and A. G. Gilman (1995). "Purification of recombinant G proteins from Sf9 cells by hexahistidine tagging of associated subunits. Characterization of alpha 12 and inhibition of adenylyl cyclase by alpha z." J Biol Chem **270**(4): 1734-41.
- Kriminski, S., C. L. Caylor, et al. (2002). "Flash-cooling and annealing of protein crystals." Acta Crystallogr D Biol Crystallogr **58**(Pt 3): 459-71.
- Kwong, P. D. a. L., Y. (1999). "Use of cryoprotectants in combination with immiscible oils for flash cooling macromolecular crystals." J Applied Cryst **32**: 102-105.
- Minor, Z. O. a. W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods in Enzymology. (New York), Academic Press. **276**: 307-326.
- Ribolidi-Tunnicliffe, A. a. H., R. (1999). "Cryocrystallography with oil - an old idea revived." J Applied Cryst **32**: 1003-1005.
- Rodgers, D. W. (1994). "Cryocrystallography." Structure **2**(12): 1135-40.
- Rubinson, K. A., J. E. Ladner, et al. (2000). "Cryosalts: suppression of ice formation in macromolecular crystallography." Acta Crystallogr D Biol Crystallogr **56** (Pt 8): 996-1001.
- Scott, J. K., S. F. Huang, et al. (2001). "Evidence that a protein-protein interaction 'hot spot' on heterotrimeric G protein betagamma subunits is used for recognition of a subclass of effectors." Embo J **20**(4): 767-76.

- Ueda, N., J. A. Iniguez-Lluhi, et al. (1994). "G protein beta gamma subunits. Simplified purification and properties of novel isoforms." J Biol Chem **269**(6): 4388-95.
- Wall, M. A., D. E. Coleman, et al. (1995). "The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2." Cell **83**(6): 1047-58.

CHAPTER THREE

Structure Determination of the $G\beta_{1\gamma_2}$ •SIGK Complex

As discussed in Chapter Two, the crystals of the $G\beta_{1\gamma_2}$ •SIGK complex were optimized to a size of $150\mu \times 50\mu \times 20\mu$. However, the crystals did not diffract well using rotating anode sources, necessitating the use of synchrotron radiation for crystal screening. At the synchrotrons, the crystals of the $G\beta_{1\gamma_2}$ •SIGK complex presented a challenge for data collection due to their sensitivity to radiation decay, even after cryoprotection. In addition, several factors relating to the quality of the raw data made data reduction and structure determination challenging. In this chapter, the basic concepts of structure determination using molecular replacement methods will be explained; in addition, the procedures undertaken to determine the structure of the $G\beta_{1\gamma_2}$ •SIGK complex will be discussed.

MATERIALS AND METHODS

Data Collection and Data Reduction

Data collection was performed at the 8.2.1 beamline at the Advanced Light Source (ALS) in Berkeley, CA using in-house software (DCS). The wavelength used was 1.0781λ (11.5 keV); the distance between the crystal and the ADSC Quantum 315 CCD detector was 250 mm. The exposure time was 30s per frame; 1° oscillation rate per frame was used. 180° degrees of rotation (180 frames) were collected on the crystal used for structure determination. HKL2000 was used at the 8.2.1 beam line to assess diffraction quality, perform initial indexing, and identify data collection

strategy based on the crystal orientation and initial point group determination (Minor 1997). The DENZO and SCALEPAK components of the HKL2000 program were utilized for data reduction (Minor 1997).

Identifying the Correct Space Group of the G $\beta_1\gamma_2$ •SIGK Crystals

To identify the correct space group of the G $\beta_1\gamma_2$ •SIGK data set, systematic absences were analyzed using the SCALEPAK module within HKL2000 (Minor 1997). In addition, all possible orthorhombic space group possibilities (P222, P222₁, P2₁2₁2, and P2₁2₁2₁) were tested in CNS, AMoRE, and PHASER to look for productive rotation:translation solutions.

Introduction to Macromolecular Crystallography: The Phase Problem

The goal of X-ray crystallographic analysis is to record the diffraction pattern from a solid crystal using an X-ray source and a detector; to use these diffraction intensities to figure out the structure of the crystal, and thus the atomic positions for each residue in the macromolecule under consideration (McPherson 2003). The elements of the diffraction spectra (the structure factors) for a crystal may be written as Equation 3-1.

$$\mathbf{F}_h = V \sum_{\substack{\text{all} \\ \text{atoms} \\ j}} \mathbf{f}_j \exp i 2\pi(\mathbf{r}_j \cdot \mathbf{h})$$

Equation 3-1. The structure factor equation. $\mathbf{F}_h = \mathbf{F}_{hkl}$ = structure factor; V = volume of the unit cell; \mathbf{f}_j , the scattering factor of all the atoms (j) in a unit cell; \mathbf{r}_j , the coordinates x_j, y_j, z_j of the j^{th} atom; \mathbf{h} , the reciprocal lattice vector.

Derivations of all equations may be found in McPherson (McPherson 2003). The consequence of this equation is that given the atomic coordinates of a crystal, we can calculate the intensities at each reciprocal lattice point in the diffraction pattern of the crystal – and the reverse must also be true. In practice, the crystallographer utilizes the diffraction data and the structure factors to calculate the electron density $\rho(\mathbf{x}, \mathbf{y}, \mathbf{z})$ for all points within the crystallographic unit cell (McPherson 2003). This equation takes the form of Equation 3-2.

$$\rho(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \frac{1}{V} \sum_{\mathbf{h}=-\infty}^{+\infty} |\mathbf{F}_{\mathbf{hkl}}| \cos[2\pi (\mathbf{h}\mathbf{x} + \mathbf{k}\mathbf{y} + \mathbf{l}\mathbf{z}) + \varphi_{\mathbf{hkl}}]$$

Equation 3-2. The electron density equation. $\mathbf{F}_{\mathbf{hkl}}$, V , and \mathbf{h} are as in Equation 3-1. For given values of $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in real space, summation is over $\mathbf{h}, \mathbf{k}, \mathbf{l}$ in reciprocal space, thereby providing the running indices of the coefficients in the series. $\varphi_{\mathbf{hkl}}$ is the phase angle in radians of $\mathbf{F}_{\mathbf{h}}$.

This equation states that given structure factors for any point $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in the unit cell and a simple trigonometric term, one can calculate directly the value of the electron density about each grid point. However, the difficulty lies in the fact that $\mathbf{F}_{\mathbf{hkl}}$ is a vector quantity, a complex number described by both amplitude and phase. The detection of x-ray diffraction can only provide the amplitudes of the structure factors, but not the phases. It is therefore necessary to supply phase information through

various methods (heavy atom diffraction, anomalous scattering, or molecular replacement) in order to overcome this issue and solve Equation 3-2.

Solving the Phase Problem using Molecular Replacement

The fundamental problem in structure determination using x-ray crystallographic data is that the calculation of an electron density map requires both the diffraction intensities collected during native data collection and phase associated with the diffraction at each reciprocal lattice point. For macromolecular crystals this phase information is calculated by using heavy atom diffraction, by anomalous diffraction from selenomethionine molecule incorporation, or by using phase information from a known structure. The last method of obtaining phase information is commonly called molecular replacement (Grosse-Kunstleve and Adams 2001). There are many useful discussions of the concepts involved in molecular replacement, including hard copy references (Brunger 1997) and web based tutorials (<http://www-structmed.cimr.cam.ac.uk/course.html>). The basic concept of molecular replacement is to model the set of structure factors that define the phases and amplitudes for each reflection, and relies on the similarity between the search model and the target structure. If the conformation of the search model and the target structure are similar enough, the atomic coordinates of the search model can be used to calculate the phase component of the structure factors of the target structure, thereby leading to an initial solution for the electron density map that will allow for further model building and refinement. A general estimate for the degree of

similarity needed is that molecular replacement will fail if the root mean square differences in atomic position between the search model and the crystal structure are greater than 1 Å for a structure of moderate resolution (Brunger 1997). The basic goal of all molecular replacement software is to orient and position the search model so that it coincides with the arrangement of the target protein in the crystal. In order for the orientation to be correct, six degrees of freedom (three rotational variables and three translational variables) must be determined. The three molecular replacement protocols described below (CNS (Brunger, Adams et al. 1998), AMoRE (Collaborative 1994; Navaza 1994), and PHASER (Storoni, McCoy et al. 2004)) differ in the way rotation and translation variables are optimized, and they score the success of their searches using different methods. In all cases, however, the resultant rotation and translation solutions are applied to the input search model, forming the initial molecular replacement solution and providing initial phase information. This information, obtained from any of the three programs outlined below, is fed into CNS in order to compute the electron density map (Equation 3-2) that will be used for model building and refinement (Brunger, Krukowski et al. 1990; Brunger 1993; Rice and Brunger 1994; Adams, Pannu et al. 1997; Read 2001). The details of map calculation are addressed after a discussion of the programs utilized for molecular replacement.

Using CNS for Molecular Replacement

CNS splits the positional search for the correct solution into three-dimensional rotation and three-dimensional translation searches (Grosse-Kunstleve and Adams 2001). Most of the benchmark molecular replacement programs use this procedure, following the original ideas of Rossmann and Blow concerning the superposition of Patterson functions to find the orientation of multiple protein chains in the asymmetric unit (Rossmann and Blow 1962). Although a full six-dimensional search is feasible with current computing power, it is unnecessary in order to obtain the correct solution. The search for angular coordinates is performed using “direct rotation” search methods in CNS (DeLano 1995). Most rotation search programs superimpose the Patterson map computed from the search model upon the observed Patterson map, and then a pattern matching procedure is used to compare the quality of the overlay (see Figure 3-1 for an overview of this process). The reciprocal space direct rotation search implemented in CNS rotates a search model relative to the crystal unit cell and then calculates the structure factors for the rotated search model in a P1 cell identical to the target crystal dimensions. The squared, normalized structure factor amplitudes from the rotated search model are compared to the observed diffraction data (Figure 3-1 and Equation 3-3). This is a computationally expensive procedure, but allows for fewer approximations related to the process of rotating two Patterson maps relative to each other (Grosse-Kunstleve and Adams 2001).

The scoring function is expressed as a linear correlation coefficient (Patterson correlation, PC) of the observed and the calculated normalized structure factor amplitudes (Equation 3-3) (DeLano 1995). PC refinement is then carried out, during which molecular dynamics is used to adjust the model, followed by computation of normalized P1 structure factors and calculation of a new correlation coefficient. This step has been shown to increase the likelihood of success in the later translation search (Adams, Pannu et al. 1997) and to increase the signal to noise ratio in the resultant translation search solutions (DeLano 1995; Grosse-Kunstleve and Adams 2001).

One of the functions implemented in CNS is a fast translation function introduced by Navaza and Vernoslova (Navaza and Vernoslova 1995). The target function for the fast translation function is the correlation coefficient in terms of the intensities of observed structure factors (Navaza and Vernoslova 1995).

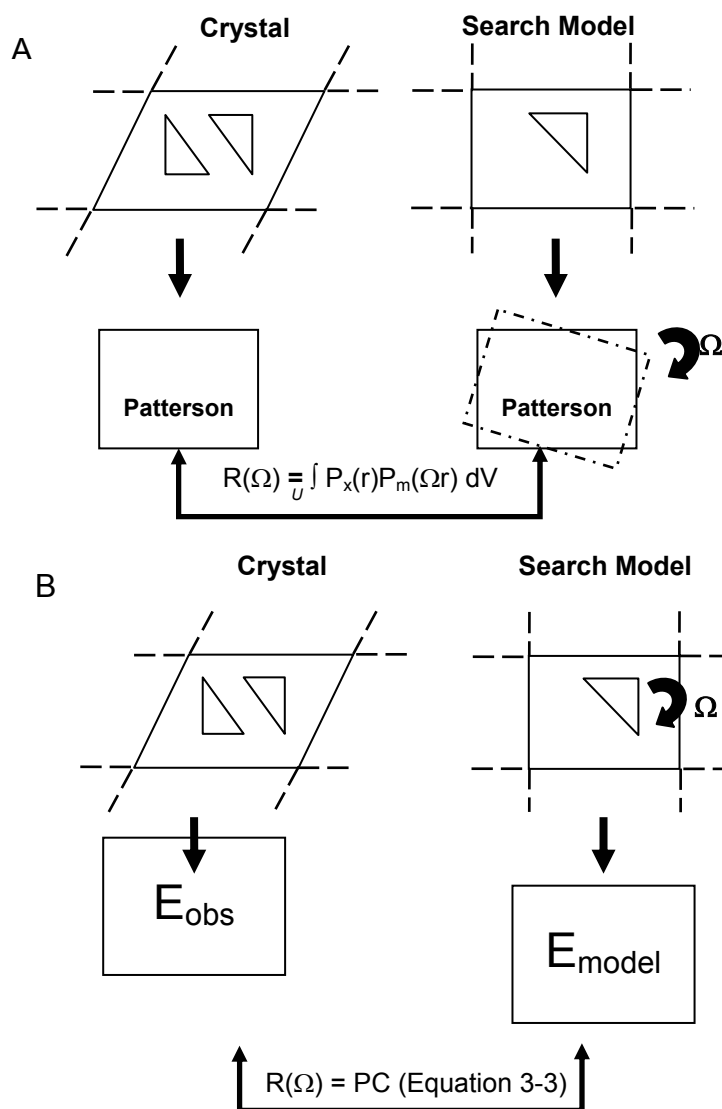


Figure 3-1. Graphical comparison of the conventional rotation function and the direct rotation function. (A) In traditional rotation functions, the Patterson map is calculated from a search model and rotated, and the value of Ω is varied to find the maximum value of $R(\Omega)$. The product function is shown; P_x = the Patterson function for the observed data; P_m = the Patterson function for the search model; r = the integration variable; U is the volume of integration, centered at the origin; V is volume. (B) In the reciprocal space rotation function implemented in CNS, the model is rotated, and then structure factors from the model and the observed data are compared using the linear correlation coefficient between squared normalized structure factors (PC, shown below). Modified from (DeLano 1995).

$$R_{\text{direct}(\Omega)} = \text{PC} = \frac{\langle |\mathbf{E}_{\text{obs}}|^2 |\mathbf{E}_{\text{m}(\Omega)}|^2 \rangle - \langle |\mathbf{E}_{\text{obs}}|^2 \rangle \langle |\mathbf{E}_{\text{m}(\Omega)}|^2 \rangle}{\{[\langle |\mathbf{E}_{\text{obs}}|^4 \rangle - \langle |\mathbf{E}_{\text{obs}}|^2 \rangle^2] \times [\langle |\mathbf{E}_{\text{m}(\Omega)}|^4 \rangle - \langle |\mathbf{E}_{\text{m}(\Omega)}|^2 \rangle^2]\}^{1/2}}$$

Equation 3-3. The target of the direct rotation function. The equation is a linear correlation coefficient between the observed normalized structure factors, \mathbf{E}_{obs} , and the calculated normalized structure factors $[\mathbf{E}_{\text{m}(\Omega)}]$. The symbols $\langle \rangle$ represent an average calculated over the observed reflections expanded into a P1 cell.

$$\text{CC}(\mathbf{x}) = \sum_{\mathbf{H}} m_{\mathbf{H}} \Delta I_{\mathbf{H}}^{\text{obs}} \Delta |\mathbf{F}_{\mathbf{H}}(\mathbf{x}) + \mathbf{F}_{\mathbf{H}}^{\text{m}}|^2 \times \left[\sum_{\mathbf{H}} m_{\mathbf{H}} (\Delta I_{\mathbf{H}}^{\text{obs}})^2 \right]^{-1/2} \times \left\{ \sum_{\mathbf{H}} m_{\mathbf{H}} [\Delta |\mathbf{F}_{\mathbf{H}}(\mathbf{x}) + \mathbf{F}_{\mathbf{H}}^{\text{m}}|^2] \right\}^{-1/2}$$

Equation 3-4. The correlation function in terms of structure factor intensities. \mathbf{H} = reciprocal space vector ($\mathbf{h}, \mathbf{k}, \mathbf{l}$); $m_{\mathbf{H}}$, the multiplicity of reflection \mathbf{H} , or the number of reciprocal vectors generated by applying the symmetry operators to \mathbf{H} ; $I_{\mathbf{H}}^{\text{obs}}$, the observed intensities = $|\mathbf{F}_{\mathbf{H}}^{\text{obs}}|^2$; $\mathbf{F}_{\mathbf{H}}(\mathbf{x})$, the calculated Fourier coefficients; $\mathbf{F}_{\mathbf{H}}^{\text{m}}$, the structure factor of the model.

Implementation of the fast translation function combined with this correlation function allows for the feasible testing for large numbers of rotation-translation solutions (Navaza and Vernoslova 1995; Grosse-Kunstleve and Adams 2001; Navaza 2001).

The combination of a fast direct rotation search, Patterson correlation refinement, and a fast translation search is more computationally intensive than the functions implemented in the AMoRE program, but may allow for small subunits or less correct search models to lead to productive solutions.

Using AMoRE for Molecular Replacement

AMoRE separates the six-dimensional search into three-dimensional rotation and translation searches for the reasons already described. The initial position of the model for the three-dimensional rotational search is placed with the center of mass at the origin and the axes of inertia oriented along the crystal symmetry axis (Navaza 2001). A fast rotation function based on the work of Crowther (Crowther 1972) is utilized to sample possible orientations of the search model. In this function the Patterson is broken down into spherical harmonics and Bessel functions, and then the rotation function is calculated using a fast Fourier transform (Equation 3-5) (Navaza 2001). The direct rotation search method implemented in CNS is at least two orders of magnitude slower in terms of CPU time than the procedure in AMoRE (DeLano 1995); however, due to the lack of approximations used in the evaluation of the Patterson correlations, the signal-to-noise ratio is significantly better for this process than the AMoRE and CNS real space rotation functions (DeLano 1995).

$$\mathcal{R}_N(R) = \int_{\Omega} P_t(r) P_s(R^{-1}r) d^3r \div \left[\int_{\Omega} P_t(r)^2 d^3r \int_{\Omega} P_s(r)^2 d^3r \right]^{1/2}$$

Equation 3-5. The normalized rotation overlap function utilized in AMoRE. The equation is in real space in terms of Patterson functions. R , the rotation matrix that superimposes the search molecule with the target crystal; P_t , the observed Patterson function; P_s , the Patterson of the search molecule. In this equation, R is in the form of a correlation coefficient and is normalized by dividing the numerator by the norms of the truncated Patterson functions. In practice, R_N is then converted into reciprocal space, and elements of spherical Bessel functions and rotation parameterization are used to convert the above equation into a form that enables the computation of R

over a large range of test angles using two dimensional fast Fourier transforms. For additional information, see (Navaza 2001).

The linear correlation coefficient between the observed and calculated amplitudes (CC_F) is used as the criterion to screen resultant orientations. Many configurations are screened, as it is the relative differences between the calculated correlation coefficients that are significant, rather than the absolute value of the correlation coefficients (Navaza 2001). After computing correlation coefficients, the top rotation solutions are fed into single-body translation searches (the number of rotation solutions tested is decided by the user). The translation search is also performed using a group of fast functions; the form of the function is determined by the user. The possibilities for measuring the overlap between translated search models and observed data currently include: (a) centered overlap; (b) Harada-Lifchitz; (c) correlation coefficient; (d) phased translation with or without contributed external phases; or (e) n -body translation with a fixed scattering contribution. The first three models utilize centered intensities ($|I_H^*| = |I_H| - \langle |I_H| \rangle$); the last two use structure factors (Navaza 2001). The output of the functions are rotation and translation coordinates, along with correlation coefficients in terms of structure factors (CC_F) and intensities (CC_I) and an R factor calculation to enable the user to determine the fit of the solution. The fast rotation and translation algorithms allow for multiple combinations of solutions to be quickly tested allow for a fast exploration of configuration space (Navaza 2001).

$$\mathbf{F}_H^{\text{cal}} = \sum_{m=1}^M \sum_{g=1}^G f_m(\mathbf{H}\mathbf{M}_g\mathbf{D}\mathbf{R}_m\mathbf{O}_m) \exp [2\pi i\mathbf{H}(\mathbf{M}_g\mathbf{T}_m+\mathbf{t}_g)]$$

Equation 3-6. The calculation of structure factors in AMoRE. The sums are over one molecule m to M independent molecules, and from g to the g^{th} symmetry operation. \mathbf{H} , the coordinates of a crystal reciprocal vector; $\mathbf{D}\mathbf{R}_m\mathbf{O}_m$, the rotation matrix \mathbf{R}_m expressed as the orthogonalizing and deorthogonalizing matrices of \mathbf{R} ; f_m , the individual molecular scattering factors; \mathbf{T}_m , the rotation vector; \mathbf{M}_g , the space group rotation matrix; \mathbf{t}_g , the space group translation vector.

Using PHASER for Molecular Replacement

PHASER builds upon the fast rotation and translation functions implemented in AMoRE. However, PHASER differs greatly from either CNS or AMoRE in its replacement of Patterson based methods with maximum likelihood functions, which work on the principle that the best model is the one that is most consistent with all of the observations (Read 2001). If the model is changed to make the observations more statistically probable, this increases the measure of its likelihood, giving it a better scoring function (Read 2001). In the case of molecular replacement, the observations can be the intensities, or after conversion, the normalized structure factors (E values) (Storoni, McCoy et al. 2004). As in the Crowther rotation function, the likelihood-enhanced fast rotation function used in PHASER is implemented using fast Fourier transforms to decrease computational time (Storoni, McCoy et al. 2004).

$$l^s(\mathbf{k}) = D^2 F_{j_r}^2(\mathbf{k}) - \langle D^2 F_{j_r}^2 \rangle$$

Equation 3-7. The likelihood-enhanced rotation function. This is a scaled and variance-weighted version of the Patterson overlap function used in the Crowther target (Storoni, McCoy et al. 2004). D is the fraction of the calculated structure factors which are correlated to the observed values; F_{j_r} corresponds to the structure factor component with unknown relative phase of the rotating molecule j_r . Although formally there are contributions $l^t(\mathbf{h})$ and $l^s(\mathbf{k})$ in the series expansion of the Wilson likelihood function, in practice only the above symmetry related contribution is needed as a target (Storoni, McCoy et al. 2004).

PHASER reports the rotation functions as log-likelihood gain scores, which can be understood as the difference between the score from a molecular replacement trial and the score from a set of random atoms (Storoni, McCoy et al. 2004). The same sort of likelihood function is then applied to the translation function, together providing a very fast and extremely powerful tool to pull out molecular replacement solutions using partial or dissimilar search models. This is due to the fact that the likelihood-enhanced targets are more sensitive to the correct orientation of the search model than the traditional Crowther functions (Storoni, McCoy et al. 2004).

In conclusion, molecular replacement succeeds in many cases, resulting in interpretable electron density. When molecular replacement fails utilizing one software suite, switching to another program may allow the user to find a rotation and translation solution due to the different methods used to find and score solutions in the different programs. In addition, changes to the initial search model should

always be considered, as small differences in loops or secondary structural elements may result in failed molecular replacement trials.

Model Refinement

Model refinement was accomplished using the CNS suite (Brunger, Krukowski et al. 1990; Rice and Brunger 1994; Adams, Pannu et al. 1997). First, an anisotropic B-factor correction was applied, and then torsional simulated annealing was performed on the model. Next, coordinate and individual B-factor minimization was done. Target sigma values on main chain and side chain bonds and angles for restrained B-factor refinement were not changed from default values until the final stages of refinement; and the weight for B-factor restraints (R_{weight}) was optimized in the final stages of model refinement. The refinement target was a maximum likelihood target using amplitudes and phase probability distribution; the W_a weight for the X-ray term was optimized in the final stages of refinement.

Map Calculation

The map calculation was accomplished utilizing the CNS program. Sigma-A weighted maps (Read 1986) were calculated using 2Fo-Fc or Fo-Fc coefficients. These maps were calculated after each round of refinement.

Model Building

Model building was accomplished using the program O (Jones, Zou et al. 1991). Initially a polyalanine model was utilized to fit the observed electron density for the SIGK peptide. Then the correct sequence was substituted for the polyalanine model, based on the identification and orientation of the electron density for a few clear side chains (Lys4, Phe6, and Tyr11). R and R_{free} were monitored after refinement of each new model, and the quality of the resultant 2Fo-Fc and Fo-Fc σ A-weighted maps were used after each model change to assess the quality of the new crystallographic model.

Model Validation

To validate the final crystallographic model, the program PROCHECK was used within the CCP4 suite in order to assess the distribution of residues in the Ramachandran plot, and other statistical parameters including main chain and side chain bond lengths and angles, and planarity of aromatic rings. In addition, the model validation portals STAN and MolProbity

(STAN: http://xray.bmc.uu.se/cgi-bin/gerard/rama_server.pl;

MolProbity: <http://kinemage.biochem.duke.edu/molprobity/index-king.html>)

were used to identify possibly ambiguous geometry and incorrect rotamer assignment, among other validation testing.

RESULTS AND DISCUSSION

Identifying the Correct Space Group for G $\beta_1\gamma_2$ •SIGK Crystals with PHASER

The initial auto-indexing results from HKL200 indicated the G $\beta_1\gamma_2$ •SIGK complex had crystallized in an orthorhombic point group (P222) with cell constants of $a = 45.462$ Å, $b = 74.446$ Å, $c = 107.922$ Å, and $\alpha = \beta = \gamma = 90^\circ$. The Matthews coefficient (V_m) (Matthews 1968) was then calculated to estimate the solvent content and the number of molecules in the unit cell. V_m is the ratio of unit cell volume to protein molecular weight and has units of Å³/Dalton:

$$V_m = \frac{\text{unit cell volume}}{M_r \times Z}$$

where M_r is the molecular weight of the protein and Z is the number of protein monomers in the unit cell. For most protein crystals in the range of 40-60 kDa, V_m is between 2.0 and 3.5 Å³/Dalton (Kantardjieff and Rupp 2003). For the G $\beta_1\gamma_2$ •SIGK crystals, V_m was 1.98 Å³/Dalton, corresponding to a solvent content of 38% for one molecule in the asymmetric unit (four molecules in the unit cell).

To identify possible screw axes, the systematic absences along each two fold axis in the crystal were analyzed after integration and scaling of the raw data. This is the standard method for assigning the space group of a crystal, although several factors can obscure systematic absences thereby leading to mis-assignment of the space

group. For this reason, the user is allowed in HKL2000 to reindex and scale the data set in all possible space group settings allowed for a particular point group, and then utilize this data in the next step of structure solution. In the case of the $G\beta_1\gamma_2$ •SIGK complex, the initial space group was assigned as $P222_1$ based on the systematic absences (Table 3-1), although the possibility that the correct space group was $P2_12_12$ was not discounted. Due to this ambiguity, all possible orthorhombic space groups ($P222_1$, $P2_12_12$, $P2_12_12_1$) were used in molecular replacement trials.

Intensities of systematic absences					
h	k	l	I	$\sigma(I)$	$I/\sigma(I)$
0	0	3	9.5	5.0	1.9*
0	0	5	5.2	5.6	0.9
0	0	7	23.6	8.6	2.7*
0	0	9	50.7	13.5	3.7*
0	0	11	53.8	13.5	4.0*
0	0	13	258.4	17.3	14.9*
0	0	15	21.1	13.5	1.6*
0	0	17	10.6	14.4	0.7
0	0	19	90.3	18.8	4.8*
0	0	21	263.7	24.2	10.9*
0	0	23	692.8	32.7	21.2*
0	0	25	2.3	30.7	0.1
0	0	27	251.8	40.9	6.2*
0	0	29	-19.3	36.4	-0.5
0	0	35	90.6	44.8	2.0*
0	0	37	13.3	52.2	0.3
0	3	0	10.8	9.9	1.1
0	5	0	18.5	15.0	1.2
0	7	0	3.0	20.0	0.2
0	9	0	70.5	27.2	2.6*
0	11	0	64.1	31.6	2.0*
0	13	0	-25.7	30.3	-0.8
0	15	0	41.5	42.0	1.0
0	17	0	-30.2	44.6	-0.7
0	19	0	74.9	69.8	1.1
0	21	0	97.6	86.3	1.1
0	23	0	558.9	146.5	3.8*
0	25	0	14.3	66.3	0.2
0	27	0	-23.6	51.0	-0.5
3	0	0	42.6	5.2	8.2*
5	0	0	25.0	6.5	3.8*
15	0	0	-1.2	14.3	-0.1

Table 3-1. Systematic absences along the three axes of the SIGK•G $\beta_{1\gamma_2}$ data set used for structure determination. I = intensities; σ = standard deviation. $I/\sigma(I)$

values greater than 1.2 are starred. A preponderance of these values indicates the lack of systematic absences along that axis.

It is symptomatic of an incorrect space group assignment for a rotation solution to be found in molecular replacement, but for translation functions to fail. This is due to the fact that all rotational search algorithms search for solutions in P1 space, and is therefore independent of the space group of the data set. The particular space group plays into the translation function solution; if the space group is incorrectly assigned, the translation functions will encounter steric clashes when expanded out from the single molecule solution, thereby leading to unproductive phase calculation and uninterpretable electron density.

In the case of the initial structure solution of the $G\beta_{1\gamma_2}$ •SIGK complex, attempts to find a correct translation function using either CNS or AMoRE failed, although both rotation algorithms seemed to find unique and well-correlated solutions. The initial search model consisted of the full structure of the $G\beta_{1\gamma_2}$ heterodimer as found in the GRK2• $G\beta_{1\gamma_2}$ complex (1OMW (Lodowski, Pitcher et al. 2003)). This model was used because it was crystallized in the presence of the wild type lipid modification to the $G\gamma_2$ molecule, and in the presence of the detergent CHAPS, making it the most similar search model to the proteins found in the SIGK• $G\beta_{1\gamma_2}$ complex. Alternate space groups were tested; P222, P222₁, and P2₁2₁2 were all screened, to no avail. Finally, a lower symmetry space group was tested, in the hope that two of the two-fold symmetries were pseudosymmetric. However, P2 and P2₁ did not yield useful

solutions either. In the end, the initial $G\beta_1\gamma_2$ •SIGK structure was solved utilizing four complexes in the asymmetric unit, in the triclinic space group P1. This configuration led to a reasonable refinement, with the final model having R and R_{free} values of 22.4% and 27.3%, respectively, and other measures of model validation showing good agreement with published structures. After having completely solved the SIGK• $G\beta_1\gamma_2$ structure, it was possible through changes to the search model (using coordinates from the $G\alpha\beta\gamma$ heterotrimer or using coordinates removing all helical structure of $G\beta$ and all of $G\gamma$) and data scaled in the correct space group to find low-signal translation functions in AMoRE which led to noisy electron density. However, for these solutions no density for the SIGK peptide could be discerned. Although the $G\beta_1\gamma_2$ •SIGK model could be successfully solved in the P1 space group, it was preferable to define a higher symmetry group for the crystal since the initial indexing information suggested an orthorhombic arrangement of molecules in the crystal. Indicative of a truly orthorhombic space group, α , β , and γ angles refined to nearly 90° in P1 ($\alpha=90.03$, $\beta = 90.03$, $\gamma = 89.82$). In addition, the NCS (non-crystallographic symmetry) operators utilized in the refinement of the P1 structure fit the rotational requirements of two-fold symmetry along all three axes, although the translational relationship between the four molecules in the asymmetric unit did not precisely fit any of the possible screw axes of the orthorhombic space groups (Figure 3-2).

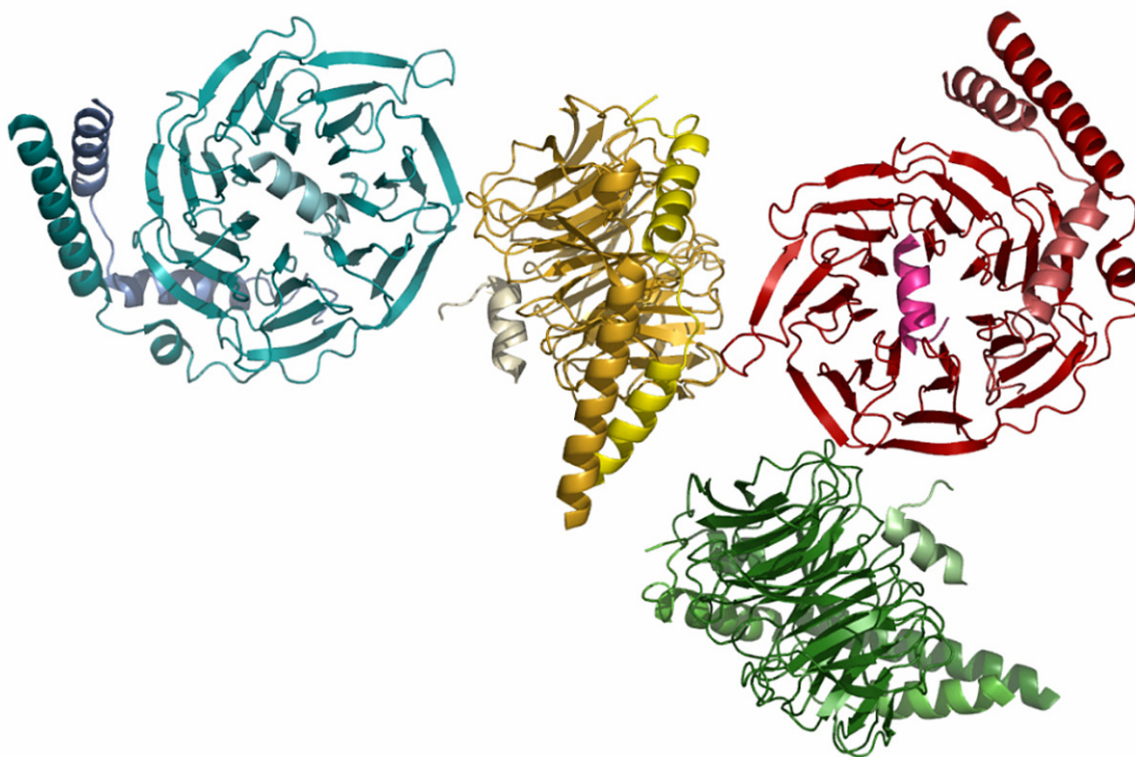


Figure 3-2. The arrangement of molecules in the P1 model. The four molecules in the asymmetric unit are shown in (A) blue, (B) yellow, (C) red, and (D) green. 2-fold rotational axes relate the four molecules to each other with RMSDs of 0.1 Å.

PHASER is a relatively new molecular replacement program, and informal reports in the crystallographic community reported a high degree of success on “difficult” structure determination. In an attempt to test the new program, the same data set was probed with the 1OMW search model which had failed to produce a reasonable rotation:translation solution using CNS and AMoRE in orthorhombic space groups. Since PHASER uses fast rotation and translation algorithms, multiple space group possibilities can be tested automatically within a single run; therefore, the program was asked to search for any possible orthorhombic solution. Within fifteen minutes PHASER had detected a single solution in space group $P2_12_12_1$ with a very high log-

likelihood score, which then refined to an R value of 27% after a single round of simulated annealing. The map obtained from this model showed clear electron density for the SIGK peptide, although this was not part of the search model. After multiple rounds of model building and refinement, the final statistics for the SIGK•G $\beta_{1\gamma_2}$ structure determined in space group P2₁2₁2₁ are listed in Table 3-2. The single molecule in the P2₁2₁2₁ asymmetric unit is not in a distinct conformation from any of the four molecules in the P1 asymmetric unit, and therefore solving the structure in the orthorhombic space group did not yield any additional insight into the SIGK•G $\beta_{1\gamma_2}$ interaction. It did, however, provide a comparison between CNS, AMoRE, and PHASER in their ability to find correct molecular replacement solutions.

A complete analysis of the G $\beta_{1\gamma_2}$ •SIGK structure is the subject of Chapter Four.

Data Collection			
Space group	P2 ₁ 2 ₁ 2 ₁	Unique Reflections	9574
Unit cell		Redundancy	3.5 (1.8)
a (Å)	45.468	Completeness (%) ¹	90.1 (56.2)
b	74.669	<I/σ> ¹	13.5 (1.6)
c	108.023	R _{sym} ^{1,2}	8.7 (41.4)
α (°)	90	Mosaicity (°)	2.3
β	90	Wilson B-factor (Å ²)	61.8
γ	90		
D _{min} (Å)	2.7		
Refinement			
Resolution (Å)	45.4 - 2.7	R.M.S. deviations	
Number of atoms ³		Bond lengths (Å)	0.006
Protein	3059	Bond angles (°)	1.3
Water	37	R.M.S. B-factors (Å ²)	
		Bonded main chain	1.29
R _{work} (%) ⁴	22.7	Bonded side chain	1.81
R _{free} (%) ⁵	28.7	Average B-factor (Å ²) ⁶	46.3

Table 3-2. Refinement statistics for the SIGK•Gβ₁γ₂ complex.

¹Numbers in parentheses correspond to the highest resolution shell, 2.8-2.7 Å.

² $R_{\text{sym}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ and $\langle I(h) \rangle$ are the i^{th} and mean measurement of the intensity of reflection h , respectively. ³ The final model contains residues 2-340 of Gβ₁ (of 340), 7-52 of Gγ₂ (of 68), 1-13 of SIGK (of 15), and 37 water molecules. ⁴ $R_{\text{work}} = \sum_h ||F_o(h)| - |F_c(h)|| / \sum_h |F_o(h)|$, where $F_o(h)$ and $F_c(h)$ are the observed and calculated structure factors, respectively. An I/σ cutoff was not used in the final calculations of R-factors. ⁵ R_{free} is the R-factor obtained for a test set of reflections consisting of a randomly selected 8% of the data. ⁶B-factors at the N-termini, including Gβ₁ residues 2-41 and Gγ₂ residues 7-13, are greater than 80 Å².

BIBLIOGRAPHY

- Adams, P. D., N. S. Pannu, et al. (1997). "Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement." Proc Natl Acad Sci U S A **94**(10): 5018-23.
- Brunger, A. T. (1993). "Assessment of phase accuracy by cross validation: the free R value. Methods and applications." Acta Crystallogr D Biol Crystallogr **49**(Pt 1): 24-36.
- Brunger, A. T. (1997). Patterson Correlation Searches and Refinement. Upton, Academic Press.
- Brunger, A. T., P. D. Adams, et al. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr **54** (Pt 5): 905-21.
- Brunger, A. T., A. Krukowski, et al. (1990). "Slow-cooling protocols for crystallographic refinement by simulated annealing." Acta Crystallogr A **46** (Pt 7): 585-93.
- Collaborative (1994). "The CCP4 suite: programs for protein crystallography." Acta Crystallographica Section D **50**(5): 760-763.
- Crowther, R. A. (1972). The Molecular Replacement Method. New York, Gordon and Breach.
- DeLano, W. L. (1995). "The direct rotation function: Patterson correlation search applied to molecular replacement." Acta Crystallogr D Biol Crystallogr **51**(Pt 5): 740-8.
- Grosse-Kunstleve, R. W. and P. D. Adams (2001). "Patterson correlation methods: a review of molecular replacement with CNS." Acta Crystallogr D Biol Crystallogr **57**(Pt 10): 1390-6.
- Jones, T. A., J.-Y. Zou, et al. (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models." Acta Crystallographica Section A **47**(2): 110-119.
- Kantardjieff, K. A. and B. Rupp (2003). "Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals." Protein Sci **12**(9): 1865-71.

- Lodowski, D. T., J. A. Pitcher, et al. (2003). "Keeping G proteins at bay: a complex between G protein-coupled receptor kinase 2 and Gbetagamma." Science **300**(5623): 1256-62.
- Matthews, B. W. (1968). "Solvent content of protein crystals." J Mol Biol **33**(2): 491-7.
- McPherson, A. (2003). Introduction to Macromolecular Crystallography. Hoboken, John Wiley and Sons, Inc.
- Minor, Z. O. a. W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods in Enzymology. (New York), Academic Press. **276**: 307-326.
- Navaza, J. (1994). "AMoRe: an automated package for molecular replacement." Acta Crystallographica Section A **50**(2): 157-163.
- Navaza, J. (2001). "Implementation of molecular replacement in AMoRe." Acta Crystallogr D Biol Crystallogr **57**(Pt 10): 1367-72.
- Navaza, J. and E. Vernoslova (1995). "On the fast translation functions for molecular replacement." Acta Crystallogr A **A51**: 445-449.
- Read, R. J. (1986). "Improved Fourier coefficients for maps using phases from partial structures with errors." Acta Crystallogr **A42**: 140-149.
- Read, R. J. (2001). "Pushing the boundaries of molecular replacement with maximum likelihood." Acta Crystallogr D Biol Crystallogr **57**(Pt 10): 1373-82.
- Rice, L. M. and A. T. Brunger (1994). "Torsion angle dynamics: reduced variable conformational sampling enhances crystallographic structure refinement." Proteins **19**(4): 277-90.
- Rossmann, M. G. and D. M. Blow (1962). "The detection of sub-units within the crystallographic asymmetric unit." Acta Crystallogr **15**: 24-31.
- Storoni, L. C., A. J. McCoy, et al. (2004). "Likelihood-enhanced fast rotation functions." Acta Crystallogr D Biol Crystallogr **60**(Pt 3): 432-8.

CHAPTER FOUR

Structural Analysis of the $G\beta_{1\gamma_2}$ •SIGK Complex

After solving the structure of the $G\beta_{1\gamma_2}$ •SIGK complex as outlined in Chapter Three, the structure was used to analyze the biological effects of the SIGK peptide in a structural context. In addition, the structure of the $G\beta_{1\gamma_2}$ •SIGK complex led to a comparison to the other crystal structures of $G\beta\gamma$ •protein complexes; although the SIGK peptide has no relationship to a known $G\beta\gamma$ target protein by sequence, it accesses the “top” face of $G\beta$ in a way very similar to other $G\beta\gamma$ binding proteins. Finally, the idea of $G\beta$ as a “hot spot” for protein interaction was analyzed, and conclusions about the molecular mechanism by which $G\beta$ interacts with diverse sequences and structures were drawn.

MATERIALS AND METHODS

Structure Comparison

Structural superpositions were performed using PyMol (<http://www.pymol.org>) using $G\beta$ residues 40-340 (which includes the β propeller structure but not the N terminal helix). $C\alpha$ atoms were used for superposition. These superpositions output a global RMSD fit for all atoms used in the alignment, used to calculate conformational flexibility measurements. All $G\beta$ complexes were compared to a structure of unbound $G\beta_{1\gamma_1}$ (1TBG) unless otherwise noted.

Sequence Alignment

Sequence alignment of mammalian G β subunits (human, mouse, and rat species) was carried out using the output from CLUSTAW, formatted using PrettyPlot (part of the EMBOSS suite, (Rice, Longden et al. 2000)). The input G β sequences were obtained from the NCBI database (<http://www.ncbi.nlm.nih.gov/Entrez/index.html>).

Calculations of Surface Accessibility, RMS Deviations, and B Factors per Atom

Programs in the CNS suite were utilized to measure surface accessibility, RMS deviations, and B factors per atom. Total surface accessibility was calculated on a per atom basis using a calculation based on Lee & Richards (Lee and Richards 1971). The probe radius was 1.4 Å. RMS deviations were calculated on a per atom basis with the criteria outlined in the section on structure comparison. B factor distributions were calculated using CNS; the average B factor was calculated for each residue, and also for the main chain and side chain atoms in each residue.

Software for Creating Molecular Graphics and other Figure Types

GRASP (Nicholls, Sharp et al. 1991) was utilized to calculate surface electrostatic potential for the G $\beta\gamma$ heterodimer. All other figures were generated using PyMOL (<http://www.pymol.org>).

Mutational Analysis of G β_1

All G β_1 mutants were made in the baculovirus transfer vector PDW 464 which biotinylates the protein *in vivo* specifically at a lysine upstream of the amino terminus of G β_1 . All mutants were generated by overlap extension PCR using standard protocols. Baculoviruses were generated via the Bac-to-Bac system following the manufacturer's instructions (Gibco). This work was done in the lab of A. V. Smrcka.

Small Scale Partial Purification of Biotinylated G $\beta_1\gamma_2$ (b- $\beta\gamma$) and G $\beta_1\gamma_2$ Mutants

Two hundred milliliter cultures of Sf9 cells were triply infected with baculovirus encoding (His) $_6$ -G α_{i1} , G γ_2 and either wild type or mutated G β_1 subunits. G $\beta_1\gamma_2$ dimers were purified 60 hours after infection using a modified method based on that of Kozasa and Gilman (Kozasa and Gilman 1995). Cell pellets were lysed in 4ml lysis buffer (50mM HEPES pH 8, 3 mM MgCl $_2$, 10 mM β -mercaptoethanol, 1mM EDTA, 100 mM NaCl, 10 μ M GDP and protease inhibitors) by freeze-thawing four times in liquid nitrogen. Membranes were pelleted by centrifugation at 100,000 $\times g$ for 15 minutes and bound proteins were extracted by resuspending the membranes in 4 ml of extraction buffer (50 mM HEPES pH 8, 3 mM MgCl $_2$, 50 mM NaCl, 10 mM β -mercaptoethanol, 10 μ M GDP, 1% cholate and protease inhibitors) and mixing at 4 °C for 1 hour. The suspension was centrifuged at 100,000 $\times g$ for 15 minutes and the supernatant was diluted five-fold with buffer A (50 mM HEPES pH8, 3 mM MgCl $_2$, 10mM β -mercaptoethanol, 100 mM NaCl, 10 μ M GDP, 0.5% Lubrol and protease

inhibitors). 800 μ L of pre-washed Ni-NTA resin (Qiagen) was added and mixed at 4 °C for 1 hour. The beads were pelleted by centrifugation and washed three times with 3 ml of buffer A containing 300 mM NaCl and 5 mM imidazole. $G\beta_1\gamma_2$ subunits were eluted from bound $G\alpha_{i1}$ by mixing the beads with 1ml elution buffer (Buffer A + 150mM NaCl, 5mM imidazole, 50 mM $MgCl_2$, 10 mM NaF, 10 μ M $AlCl_3$ and 1% cholate) at room temperature for 1 hour. Concentrations of eluted b- $\beta\gamma$ dimers were determined by comparing to a standard curve of fully purified 100% biotinylated $G\beta_1\gamma_2$. Proteins were separated by SDS-PAGE, transferred to nitrocellulose and probed with HRP-neutravidin (Pierce). Membranes were exposed to developing reagents and the amount of chemiluminescence was measured using an Epi-Chem II Darkroom system from UVP Bioimaging Systems. The amount of biotinylated $\beta\gamma$ was determined by comparing to the standard curve. The concentration of each mutant was determined from at least two separate gels. This work was done in the lab of A. V. Smrcka.

Phage ELISA

Phage ELISA assays used to assess peptide binding to wild-type and mutant $G\beta\gamma$ subunits were performed as described. Briefly, 1 pmol of b- $\beta\gamma$ subunits per well were immobilized on a streptavidin coated 96 well plate. Immobilized b- $\beta\gamma$ subunits were incubated with 1×10^{10} phage 3.14 (f88 phage displaying the peptide sequence SIGKAFKILGYPDYD) or an equal number of f88 phage as a control. The presence of phage was determined by adding 1:5000 dilution of HRP conjugated anti-M13

antibody (Pharmacia) and measuring the absorbance at 405 nm after the addition of ABTS to elicit a colorimetric reaction. Signals obtained with partially purified b- $\beta\gamma$ subunits were similar to signals obtained from fully purified b- $\beta\gamma$ subunits (data not shown). This work was done in the lab of A. V. Smrcka.

Determination of $G\alpha_{i1}\cdot G\beta_1\gamma_2$ Interactions by Flow Cytometry

Blocking of $G\alpha_i\cdot G\beta_1\gamma_2$ binding was assessed by simultaneously adding 200 pM FITC- $G\alpha_i$ with or without SIGK to 50 pM immobilized b- $\beta_1\gamma_2$ and measuring the amount of FITC- $G\alpha_i$ bound to the beads by flow cytometry as previously described (Sarvazyan, Remmers et al. 1998; Goubaeva, Ghosh et al. 2003). This work was done in the lab of A. V. Smrcka.

RESULTS AND DISCUSSION

Overall Structure of the $G\beta_1\gamma_2\cdot$ SIGK Complex and Mutational Analysis of the SIGK Interface

In the following text, amino acid identifiers prefixed with “s” refer to SIGK residues and otherwise to $G\alpha_{i1}$ residues. $G\beta_1$ is a β -propeller composed of seven four-stranded β -sheets (“blades”) and an N terminal extended helix that interacts extensively with $G\gamma_2$ (Figure 4-1). Each sheet is composed of WD-40 repeats connected by loops of variable length. Residues 2-340 are included in the model. B

factors throughout the core of $G\beta_1$ are less than 40 \AA^2 . B factors $> 60 \text{ \AA}^2$ are found in three loop regions: Lys127 - Ser136 in blade two, Arg214 – Met217 in blade four, and Ser265 - Ile269 in the loop connecting blades six and seven. $G\gamma_2$ forms a helix with a kink made by residues Asn24 - Lys29 and a coil region beginning at residue His44 (Figure 4-1). The average B factor within the $G\gamma_2$ molecule is 44 \AA^2 . No electron density is observed for the N terminal seven residues and the C terminal sixteen residues of $G\gamma_2$ or the prenyl lipid modification at the C terminus of $G\gamma_2$.

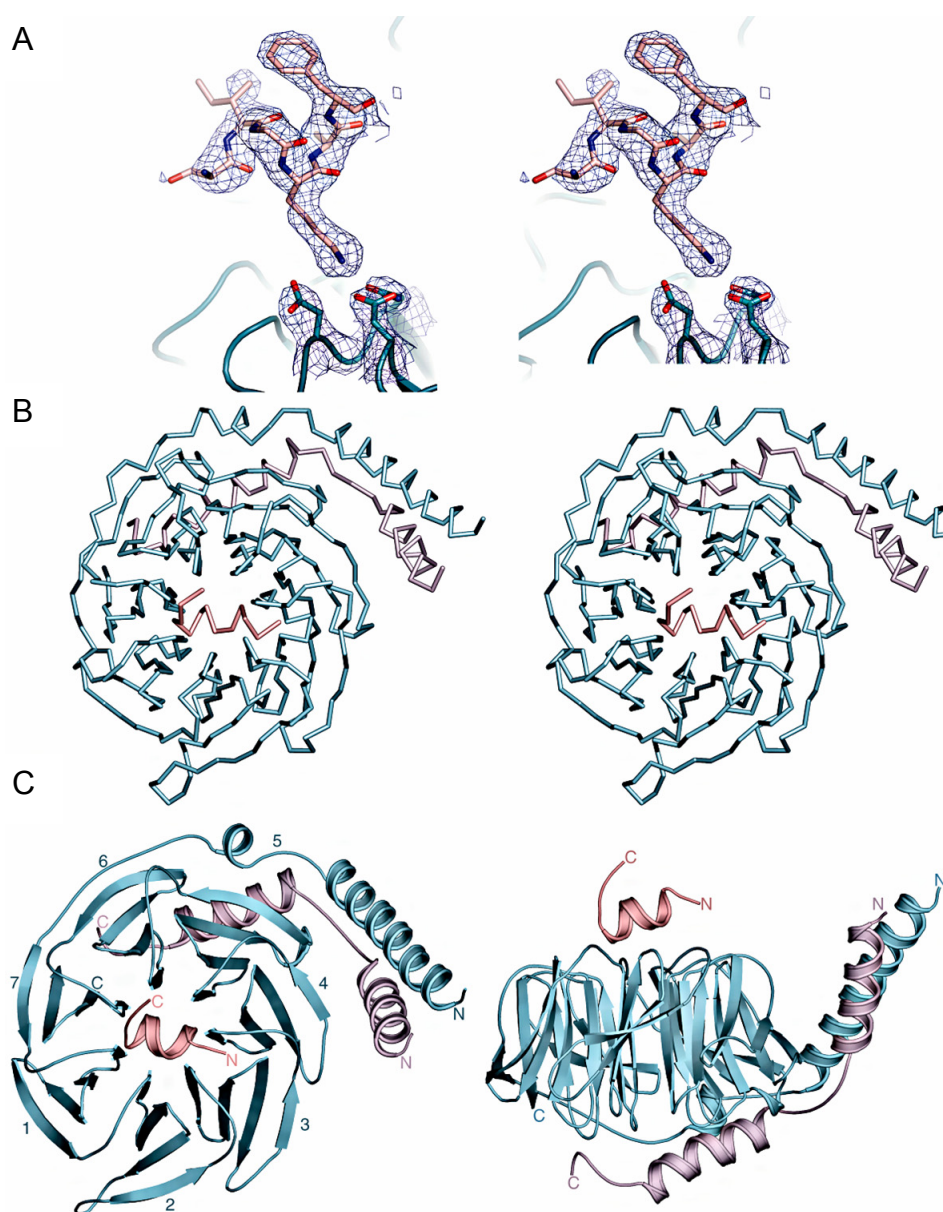


Figure 4-1. Structure of the peptide SIGK bound to the Gβ₁γ₂ heterodimer. (A) Stereo view of representative σ_A -weighted annealed 2Fo-Fc electron density, contoured at 1σ . The region shown is centered on the peptide N terminal binding site. SIGK peptide is shown in pink, Gβ₁ in blue. (B) Stereo view of the Cα trace of the Gβ₁γ₂•SIGK complex. SIGK is colored pink; Gβ₁, blue; and Gγ₂, purple. (C) Two views of the Gβ₁γ₂•SIGK complex, shown in ribbon representation. The two representations are related by -90° . Coloring is the same as in (B). The N and C termini of SIGK, Gβ₁, and Gγ₂ are labeled. On the left panel, the seven blades of Gβ₁ are labeled following the convention of Wall et al. (Wall, Coleman et al. 1995; Wall, Posner et al. 1998).

SIGK forms an α -helical structure broken by a glycine at position 10 (Figure 4-1). The C terminal three residues form an extended structure that stretches away from the $G\beta_1$ molecule and is supported by crystal contacts between sPro12 and sAsp13 with Thr47 and Lys337 from a symmetry-related $G\beta_1$ molecule. The B factors for the N- (sSer1, sIle2) and C terminal (sGly10-sAsp13) residues of SIGK are greater than 50 \AA^2 ; all other residues are between 30-50 \AA^2 . The electron density for the main chain atoms in residues 1-13 are well-defined; two SIGK side chains that do not contact $G\beta_1$ (sLys7 and sAsp13) are disordered, and sIle2 is partially ordered. The peptide binds across the “top” face of $G\beta_1$ (Figure 4-1) and buries 970 \AA^2 total solvent-accessible surface area. The peptide makes no contact with the $G\gamma_2$ subunit, which is bound to the “bottom” surface of the $G\beta_1$ torus. The SIGK contact surface on $G\beta_1$ can be separated into two regions: an acidic region on $G\beta_1$ that interacts with the N terminus of the peptide, and a largely nonpolar region that interacts with the C terminus of the peptide (Figure 4-2).

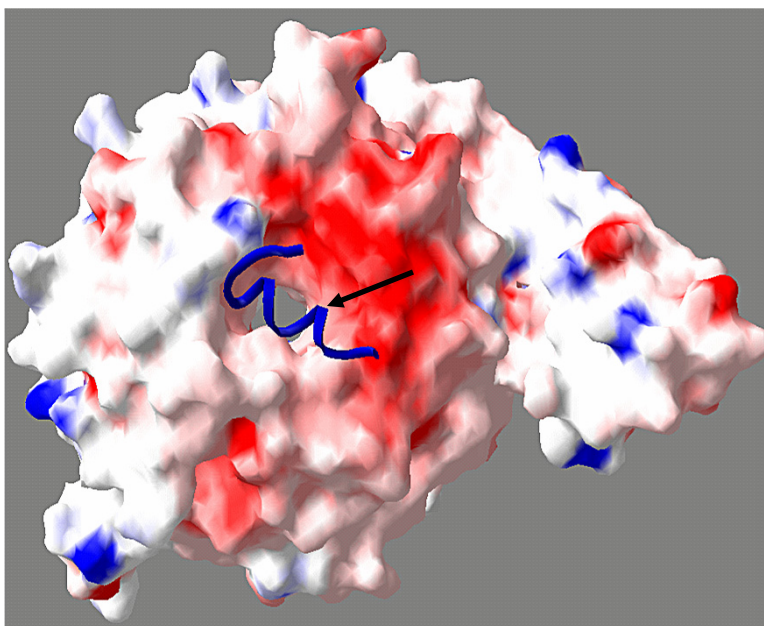


Figure 4-2. Electrostatic potential of $G\beta_1\gamma_2$. $G\beta_1\gamma_2$ is shown as a molecular surface color-coded according to electrostatic potential as calculated using GRASP (Nicholls, Sharp et al. 1991). The range of values is from -15 kT (red) to +15 kT (blue), where k is the Boltzmann constant and T is the temperature in Kelvin. SIGK peptide is shown as a $C\alpha$ trace. The negatively charged surface surrounding the N-terminal region of the peptide and sLys4 (arrow) stands in contrast to the relatively uncharged surface where the C terminal residues of SIGK bind.

In total, thirteen $G\beta_1$ residues directly contact SIGK, contributed by six of the seven blades of the β -propeller (Figure 4-3 and Table 4-1). The N terminal binding surface is centered on an electrostatic interaction in which sLys4 projects into a negatively charged binding pocket on $G\beta_1\gamma_2$ where it forms hydrogen bonded or charge interactions with Asp228, Asn230, and Asp246 (Figure 4-4). A hydrogen bond between the carbonyl oxygen of Asp228 and the main chain nitrogen of Asp246 helps stabilize the three acidic residues on $G\beta_1$. Met188 participates in van der Waals interactions with the alkyl chain of sLys4, and Asp186 forms a polar

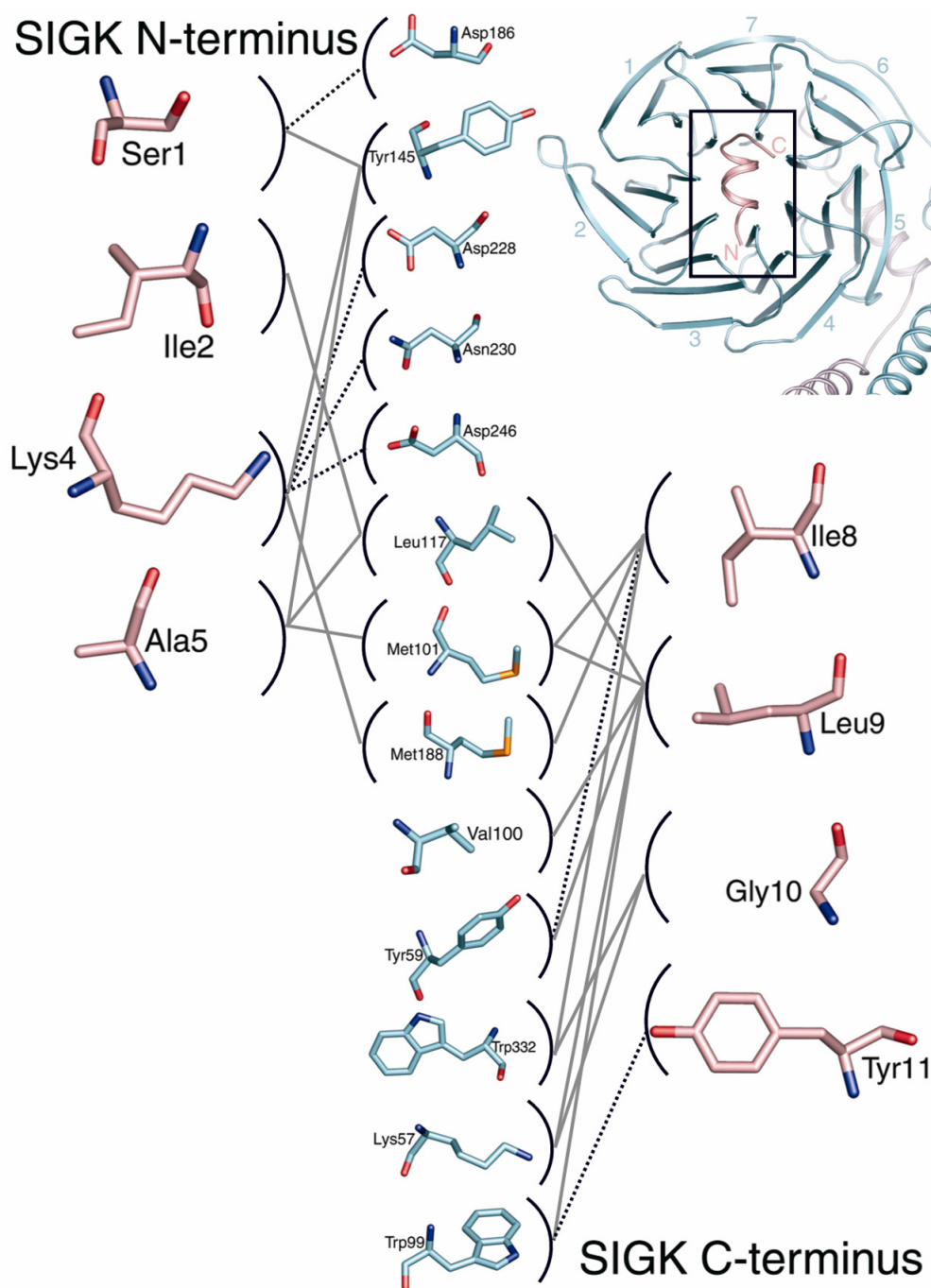


Figure 4-3. Graphical representation of the contacts between $G\beta_1$ and SIGK peptide. Residues from $G\beta_1$ are shown in blue in the middle column, flanked on the left and right columns by N and C terminal residues of the SIGK peptide, respectively. Polar contacts are represented by black dashed lines; nonpolar interactions are shown in grey unbroken lines.

Contacts at the G β_1 : SIGK Interface					
G β_1 interacting residue		SIGK interacting residue		Distance (Å)	Type of Interaction
Lys57	C ϵ	Leu9	O	3.35	Nonpolar
	C ϵ	Gly10	C α	3.99	Nonpolar
Tyr59	OH	Leu9	O	2.66	Polar
	C ϵ	Ile8	O	3.87	Nonpolar
Trp99	N ϵ 1	Tyr11	OH	2.81	Polar
	C δ 1	Leu9	C δ 2	3.59	Nonpolar
Val100	O	Leu9	C δ 2	3.75	Nonpolar
Met101	C ϵ	Ile8	C γ 2	3.46	Nonpolar
	C ϵ	Ala5	O	3.52	Nonpolar
	C ϵ	Leu9	C δ 2	3.54	Nonpolar
Leu117	C δ 1	Ile2	C γ 2	3.46	Nonpolar
	C δ 2	Ala5	C β	3.68	Nonpolar
	C δ 2	Leu9	C δ 1	3.80	Nonpolar
Tyr145	C ϵ 2	Ser1	O	3.19	Nonpolar
	OH	Lys4	C γ	3.45	Nonpolar
	C δ 2	Ala5	C β	3.81	Nonpolar
Asp186	O δ 2	Ser1	O	3.03	Polar
Met188	C ϵ	Ile8	C δ 1	3.31	Nonpolar
	C ϵ	Lys4	C ϵ	3.48	Nonpolar
Asp228	O δ 2	Lys4	N ζ	3.23	Polar
Asn230	N δ 2	Lys4	N ζ	2.82	Polar
Asp246	O δ 2	Lys4	N ζ	3.05	Polar
Trp332	C ζ 2	Ile8	O	3.12	Nonpolar
	CH2	Gly10	C α	3.57	Nonpolar

Table 4-1. Description of the SIGK binding site on G β_1 . Residues on G β_1 that are within 4Å of SIGK are listed.

contact with the carbonyl oxygen of sSer1 and also makes a hydrogen bond to the amide of Cys204. Additionally, Tyr145 forms van der Waals interactions with the main chain oxygen of sSer1, the sLys4 side chain, and the C β atom of sAla5, and forms a hydrogen bond with the nearby amide of Gly162. The side chain of Leu117 is within van der Waals contact distances of the side chains of sIle2 and sAla5. Together, these nine G β_1 residues form a surface that tethers SIGK to G β_1 using charged and nonpolar interactions.

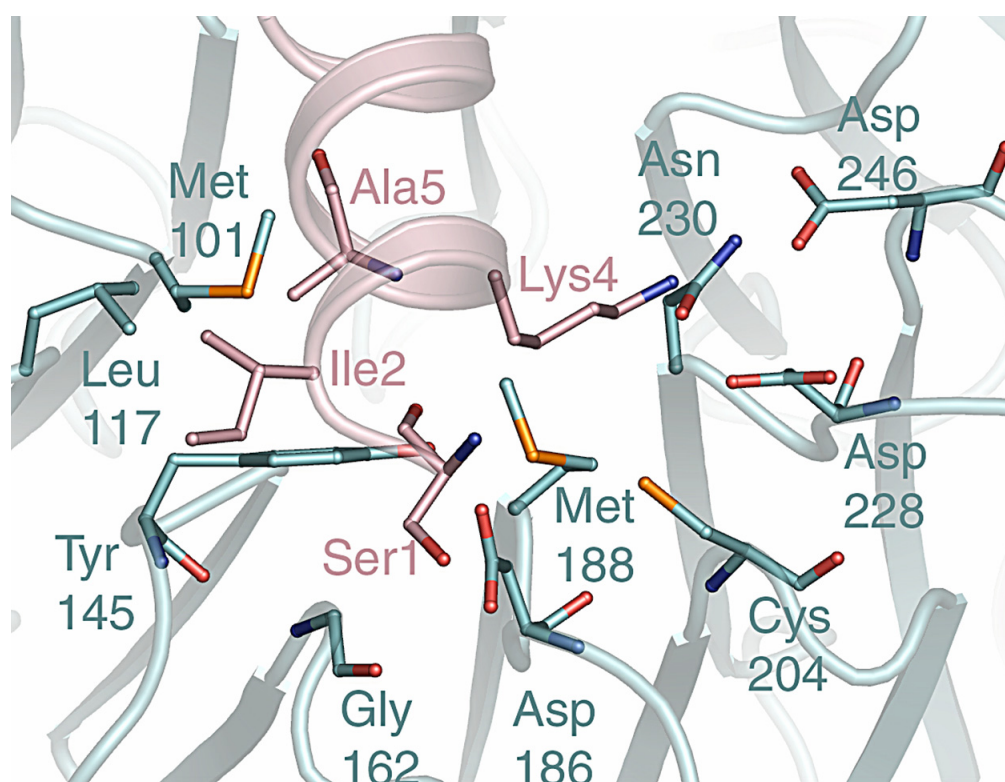


Figure 4-4. The N terminal SIGK interface with G β_1 . sLys4 of SIGK (pink) points into a highly acidic pocket on G β_1 (teal) centered around Asp228, Asn230, and Asp246 of G β_1 . Oxygen, nitrogen and sulfur atoms are colored red, blue and orange, respectively. Residues that form main chain interactions are represented with their main chain amide and oxygen visible.

Published mutational analysis of Sirk and SIGK peptides can be interpreted in the context of the SIGK•Gβ₁γ₂ structure (Scott, Huang et al. 2001; Goubaeva, Ghosh et al. 2003). Wild-type Sirk peptide inhibits the activation of PLC β₂ by Gβ₁γ₂ with an IC₅₀ of 5μM. Substitution of sLys4 to alanine in the Sirk peptide lowers the IC₅₀ of the peptide 12-fold, and mutation of sAla5 to glycine lowers the IC₅₀ by 13-fold. Mutation of sIle2 to alanine reduces IC₅₀ of the peptide by 4-fold, and mutation of sSer1 to alanine has no effect on IC₅₀ (Scott, Huang et al. 2001). The SIGK•Gβ₁γ₂ structure indicates that the main chain of sSer1 and the side chains of sIle2, sLys4, and sAla5 are contacting multiple residues on Gβ, thereby explaining this mutational data.

To measure the contribution of the Gβ₁ residues observed at the Gβ₁γ₂•SIGK interface to the binding energy for the complex, two approaches were utilized. First, an ELISA assay was used to measure binding of immobilized Gβ₁γ₂ subunits to phage displaying the SIGK sequence (Figure 4-5). The ELISA binding data were then correlated with IC₅₀ values for SIGK as a competitor of Gβ₁γ₂ association with Gα_{i1} (Figure 4-5). Both assays were then carried out with Gβ₁γ₂ heterodimers containing mutations in the Gβ₁ subunit. In the N terminal binding surface, mutation of Gβ₁ Asn230 to alanine decreases the affinity of Gβ₁γ₂ for peptide 10-fold (Figure 4-5). Single mutation of Gβ₁ residues Asp186, Met188, Tyr145, and Leu117 to alanine also resulted in Gβ₁γ₂ dimers with drastically decreased affinity for SIGK

(Figure 4-5). $G\beta_1$ mutants in which either Asp228 or Asp246 were substituted with alanine did not dimerize with $G\gamma_2$ and so were not analyzed in this study.

However, a mutant where Asp228 was substituted with serine caused only a slight loss in binding affinity for SIGK peptide (Figure 4-5). Clearly, many of the $G\beta_1$ residues which comprise the N terminal SIGK binding interface are contributing strongly to the energy of binding.

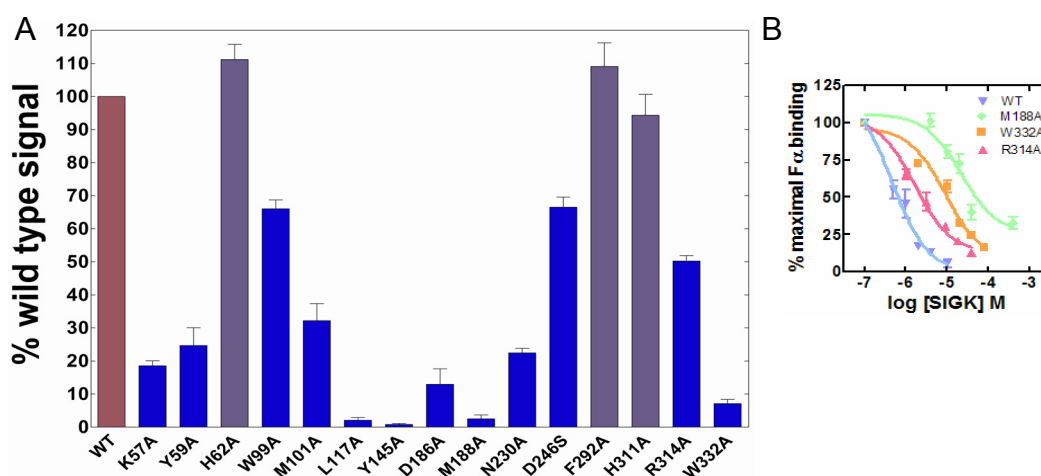


Figure 4-5. Binding of SIGK peptide to $G\beta_1\gamma_2$ mutants. (A) Amino acids that contact the SIGK peptide were individually mutated to alanine and binding to peptide was assayed using a phage ELISA. Data shown is the mean \pm SD of triplicate determinations from three independent experiments. Blue bars represent mutants for which binding is statistically different from wild type ($p < 0.001$ as determined by a one-way ANOVA followed by a Bonferroni's post test). Grey bars represent data statistically identical to wild type. (B) SIGK competition for FITC- $G\alpha_{i1}\beta_1\gamma_2$ interactions with representative $G\beta_1$ subunit mutants. Data is shown as the mean of triplicate determinants \pm standard deviation of a representative experiment. The experiment was repeated two (Met188A) or three (wild type, Arg314A, Trp332A) times with similar results. Comparison of the two assays over a selection of mutants that spanned the range of SIGK binding affinities indicates that a 50% loss of binding translates into a five-fold increase in IC_{50} , a 75% loss of binding corresponds to a 10-fold increase, a 90% loss is a 20-fold shift and a 98% loss is a 50 fold shift. The IC_{50} values are as follows: wild type = $0.47 \mu M$, Arg314A = $1.5 \mu M$, Trp332A = $9 \mu M$, Met188A = $22 \mu M$. This work was done in the lab of A. V. Smrcka.

The second area of binding involves most of the C terminal residues of SIGK (sAla5 – sGly11), which pack against a largely hydrophobic pocket on G β ₁ (Figure 4-6). This pocket extends 11Å from Trp332 on blade seven to Met188 in blade two. Eight G β ₁ residues are in direct contact with the C terminal surface of SIGK, and two more G β ₁ residues support the residues directly involved in the SIGK interaction. Met188, which interacts with sLys4 in the N terminal interface, is also within contact distance of the side chain of sLeu8 (see Figure 4-4 and Figure 4-6). SIGK residues sAla5, sLeu8 and sLeu9 are complimented by van der Waals interactions with Leu117, Met101, Trp99, Tyr59 and the alkyl chain of Lys57. The main chain oxygen of Val100 interacts with the side chain of sLeu9. The indole imine of Trp99 forms a hydrogen bond with the hydroxyl group of sTyr11 and the side chain of Trp332 makes contact with the main chain oxygen of sIle8 and the C α of sGly10. The side chains of Lys57 and Arg314 are positioned on either side of Trp332 and support its orientation in the binding site. Arg314 also forms a hydrogen bond with Trp332, and Lys57 with the nitrogen of Gln75, further stabilizing this interaction surface on G β ₁ (Figure 4-6). Data from alanine scanning of the peptide validate these structural observations (Scott, Huang et al. 2001; Goubaeva, Ghosh et al. 2003). Mutation of sIle8, sLeu9 or sGly10 to alanine increases the IC₅₀ for inhibition of PLC activation by 40-fold (5 μ M to 200 μ M), 60-fold and 12-fold, respectively (Scott, Huang et al. 2001). The same mutation of sLeu9 also blocks the ability of SIK to enhance ERK1/2 phosphorylation in RASM cells (Goubaeva, Ghosh et al. 2003).

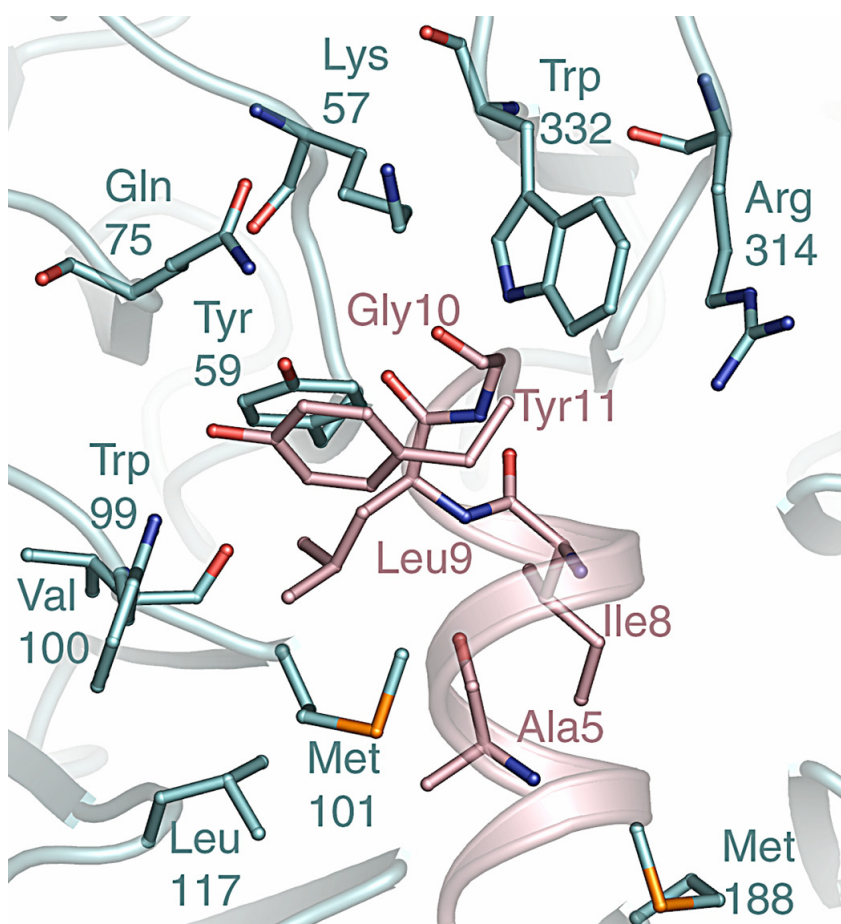


Figure 4-6. The C terminal SIGK interface with G β_1 . Residues sAla5 - sTyr11 from SIGK (pink) form an extensive interface with the mainly hydrophobic top surface of G β_1 (teal). Note that Met188 and Met101 are also involved in the N terminal interface and are shown in Figure 4-4.

Mutation of all the amino acids in G β_1 that constitute the SIGK C terminal binding surface that were tested cause a loss in affinity for the SIGK peptide, although to different extents (Figure 4-5 and Figure 4-7). Mutation of Leu117, Met188, or Trp332 to alanine nearly abrogates SIKR binding; mutants of Lys57, Tyr59, Met101, and Arg314 have more modest effects (Figure 4-5 and Figure 4-7). The Trp99

mutation results in a 4-fold decrease in affinity. A summary of all the $G\beta_1$ mutations made in the current study and their effects on SIGK binding affinity is represented in Figure 4-7. Considering all of the data for the N terminal and C terminal SIGK binding interfaces, seven of the fifteen residues of the SIGK peptide and ten of the twelve $G\beta$ residues tested contribute significant binding energy to the interface, in good correlation with the structural model.

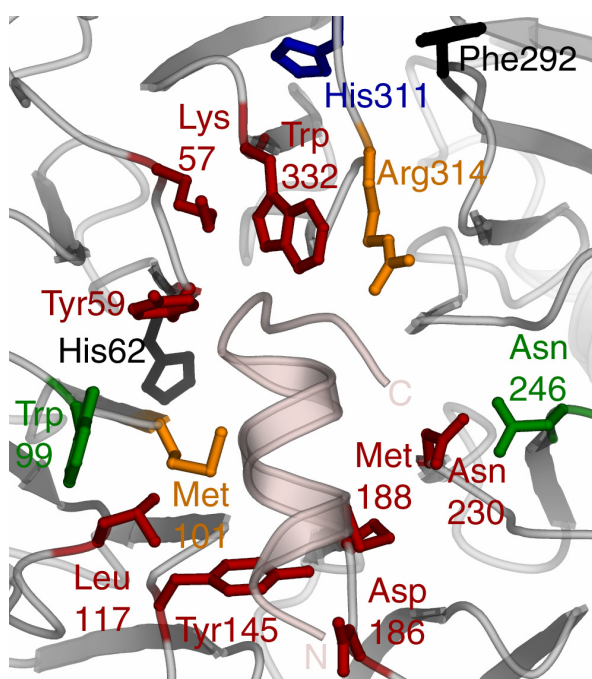


Figure 4-7. Analysis of the common protein interaction surface of $G\beta_1$. Residues mutated in the present study are shown in relation to the SIGK peptide. The N- and C- termini of the SIGK peptide (pink) are labeled. Residues of $G\alpha_1$ that, when mutated to alanine, result in 75 - 100% loss in affinity for SIGK are colored red; a 50 - 75% loss, orange; a 25 - 50% loss, green; a 0 - 25% loss, blue; no effect, black. The Asn246 mutation is to serine, not alanine.

Comparison of the $G\beta_{1\gamma_2}$ •SIGK Complex to Unbound $G\beta_{1\gamma_1}$

The binding surface of $G\beta_1$ in the $G\beta_{1\gamma_2}$ •SIGK complex is not significantly changed upon SIGK binding. The RMSD between the core residues of $G\beta_1$ in the $G\beta_{1\gamma_2}$ •SIGK complex and that in the uncomplexed $G\beta_{1\gamma_1}$ heterodimer (1TBG (Sondek, Bohm et al. 1996); Val40-Asn340, C α only) is 0.88 Å. However, the side chains of Trp99, Tyr59, Asp228, Leu117 and Met101 rotate to accommodate SIGK such that atoms within these residues undergo maximum displacements of 4.0Å, 3.6Å, 2.9Å, 2.8Å and 2.3Å, respectively, relative to their positions in uncomplexed $G\beta_1$. The B-factors for residues in the SIGK binding surface are close to the overall average for the complex. However, the B-factor for Trp99 is reduced two-fold upon binding to SIGK, as indicated by comparison of normalized B factors of the respective structures. In this analysis, there are no large conformational changes or disorder to order transitions in $G\beta$ upon SIGK binding. In addition, the SIGK• $G\beta_{1\gamma_2}$ complex may be compared to those of five $G\beta_{1\gamma_2}$ complexes with protein targets: the $G\beta_{1\gamma_2}$ • $G\alpha_{i1}$ heterotrimer (1GG2) (Wall, Coleman et al. 1995; Wall, Posner et al. 1998) and the $G\beta_{1\gamma_1}$ • $G\alpha_{vi}$ heterotrimer (1GOT) (Lambright, Sondek et al. 1996), the $G\beta_{1\gamma_1}$ •phosducin complex (1AOR and 2TRC) (Gaudet, Bohm et al. 1996; Loew, Ho et al. 1998), and the $G\beta_{1\gamma_2}$ •GRK2 complex (1OMW) (Lodowski, Pitcher et al. 2003). Superposition of the $G\beta_{1\gamma_2}$ •SIGK complex with each of these structures yields average RMS deviations for $G\beta_1$ residues 40-340 of less than 1.0 Å (C α only). The $G\beta\gamma$ heterodimer does not undergo significant structural rearrangement in order to

bind any of the protein targets seen to date, nor does it in the $G\beta_{1\gamma_2}$ •SIGK structure (Figure 4-8).

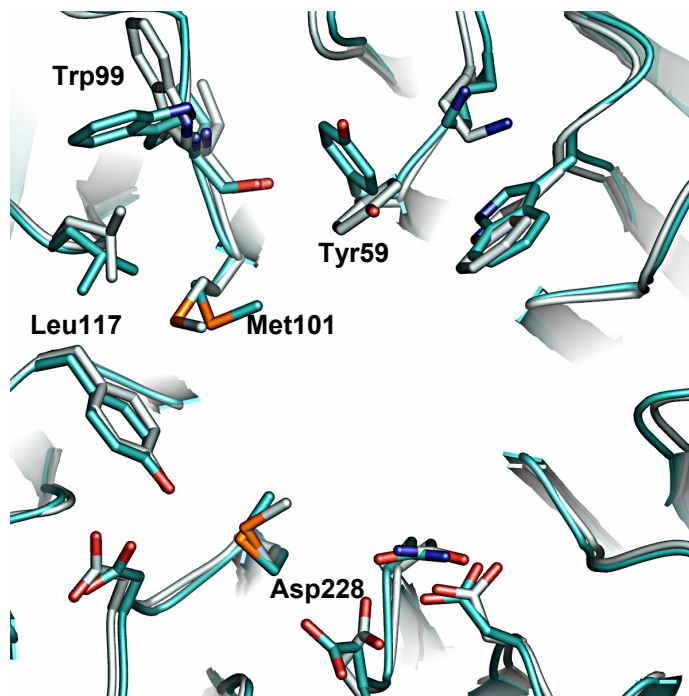


Figure 4-8. The structural changes in $G\beta$ induced by SIGK binding.

Superposition of the unbound $G\beta_1$ structure (1TBG) and $G\beta_1$ as found in the $G\beta_{1\gamma_2}$ •SIGK structure was done in PyMOL using residues 40-340. Unbound $G\beta_1$ is shown in white and $G\beta_1$ from the SIGK complex in cyan. Residues in the SIGK binding site are shown with side chains in stick representation; all other residues are shown as $C\alpha$ only. Residues with significant side chain deviation relative to the mean are labeled.

Residues in the $G\beta_1$ Binding Site Contact the SIGK Peptide and Signaling Protein

Targets

The SIGK peptide interacts with residues of $G\beta_1$ that are utilized by several $G\beta\gamma$ binding proteins. For example, Lys57, Tyr59, Trp99, Met101, Leu117, Tyr145,

Met188, Asp246, and Trp332 from $G\beta_1$ are involved in contacts with the GRK2 PH domain in the crystal structure of the $G\beta_1\gamma_2$ •GRK2 complex, and all of these residues of $G\beta_1$ are involved in SIGK contacts as well (Table 4-2). This is in spite of the fact that the secondary structures of the PH domain that contact $G\beta_1$ (the RH-PH loop, the α CT region, and β 4 strand) are completely dissimilar to the purely helical SIGK peptide (Figure 4-9) (Lodowski, Pitcher et al. 2003).

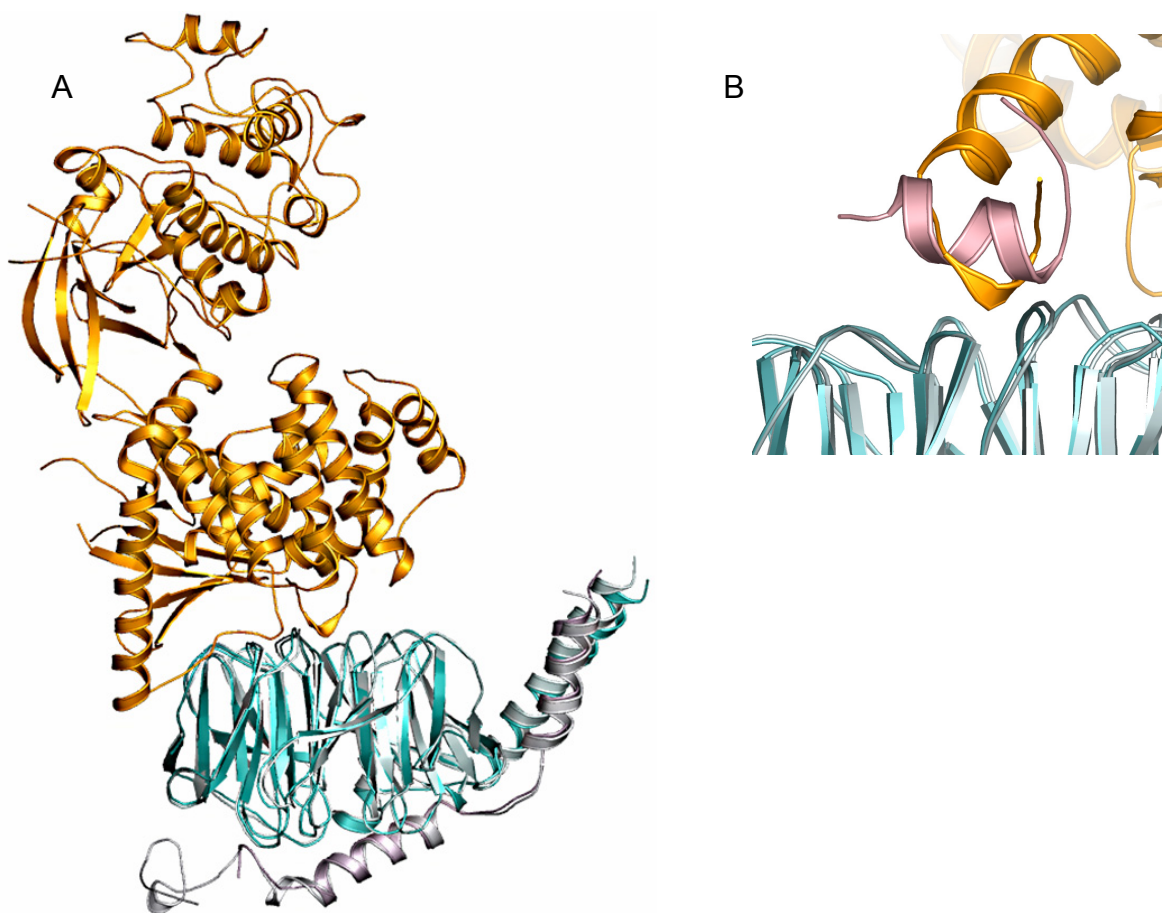


Figure 4-9. Superposition of the GRK2•Gβγ complex. In (A), the full structure of GRK2 (orange) is shown in complex with $G\beta_1\gamma_2$ ($G\beta_1$ is teal, $G\gamma_2$ is lilac). The complex is superimposed onto $G\beta_1\gamma_1$ (1TBG). In (B), the GRK2• $G\beta_1\gamma_2$ complex is superimposed onto the SIGK• $G\beta_1\gamma_2$ complex, and a view of the SIGK interface is

shown. In (B), GRK2 is orange, $G\beta_1$ from 1OMW is white, $G\beta_1$ from 1XHM is teal, and SIGK is in pink.

This theme is recapitulated in the complex of $G\beta_1$ with phosducin (Loew, Ho et al. 1998) where a common subset of $G\beta_1$ residues contacts a binding partner with different secondary structure from GRK2. All but one residue of the SIGK binding interface on $G\beta_1$ is also a contact residue for phosducin, with Val100 being the only exception. Phosducin binds the top face of $G\beta$ using secondary elements in a completely different orientation than SIGK (Figure 4-10).

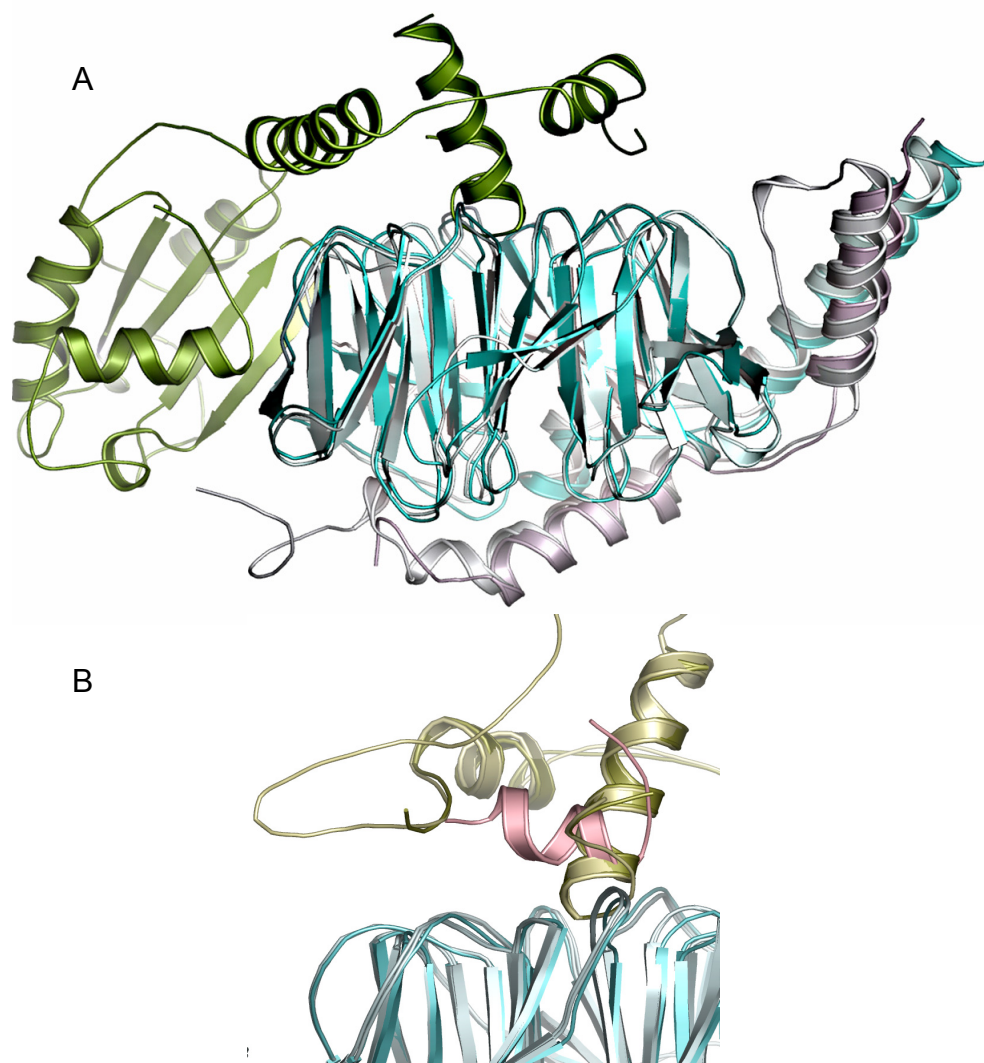


Figure 4-10. Superposition of the phosducin•Gβγ complex. The coloring scheme is roughly the same as in Figure 4-9, except phosducin is in green in (A) and olive green in (B).

Notably, the switch II region of $G\alpha_{i1}$ forms an α -helix that is bound in almost the same orientation as the SIGK peptide (Figure 4-11). However, switch II of $G\alpha_{i1}$ has no sequence similarity to the SIGK peptide, although it contains a lysine (Lys210) which is oriented in almost the same position as sLys4 (Figure 4-11) (Wall, Posner et al. 1998).

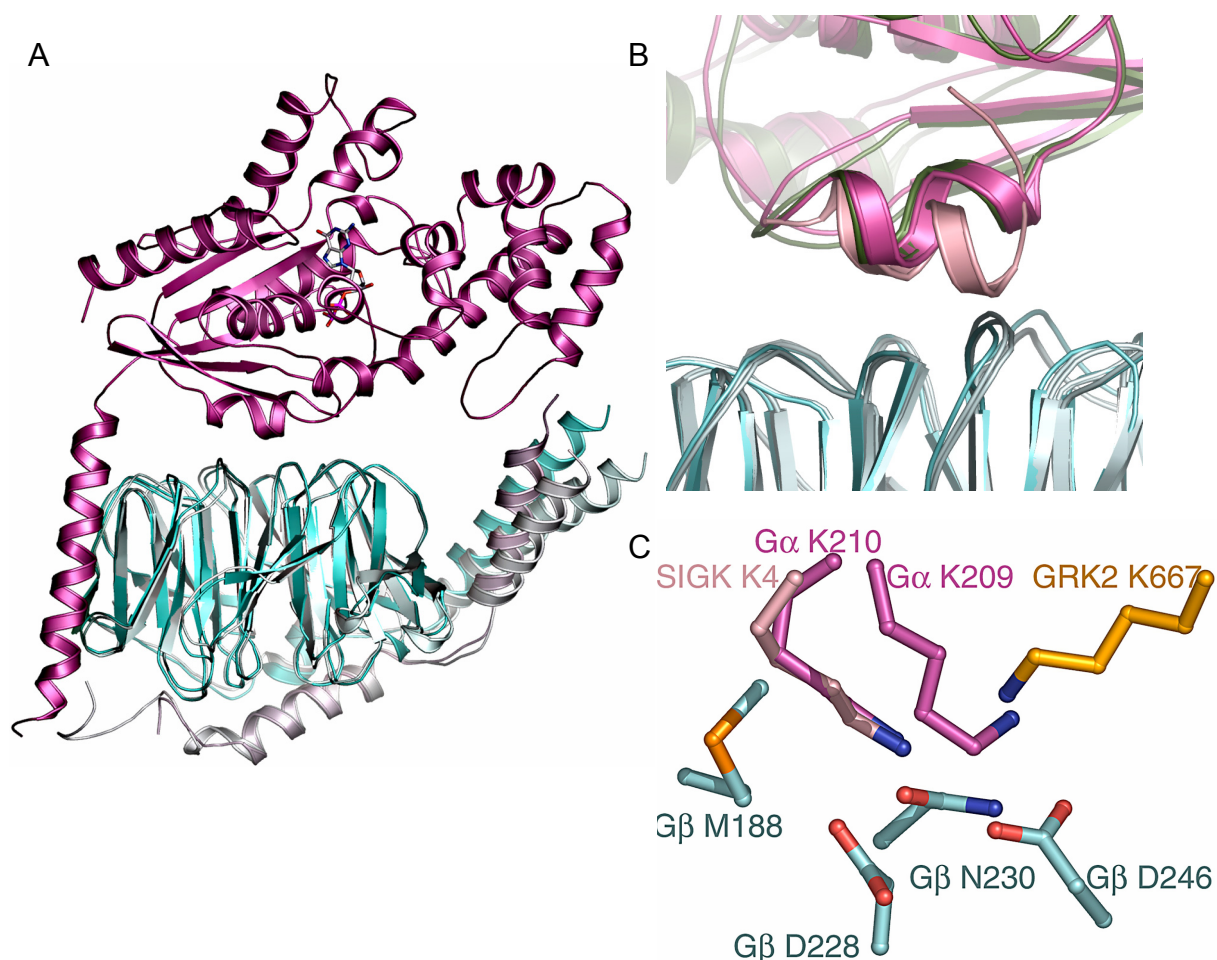


Figure 4-11. Superposition of the $G\alpha\beta\gamma$ heterotrimer. The coloring scheme is similar to that in Figure 4-9, except that $G\alpha_{i1}$ is in dark pink and $G\alpha_t$ is in olive green. In (C), side chains from $G\beta_1$ (teal), SIGK (light pink), $G\alpha_{i1}$ (dark pink), and GRK2 (orange) are shown in the N terminal SIGK binding interface.

The top face of G β is able to bind disparate sequences and structures without undergoing large conformational changes (Figures 4-8 through 4-11). However, the same residues of G β are utilized in the top face in order to accomplish this binding. If mutational data for G $\beta\gamma$ targets PLC β 2, adenylyl cyclase, and GIRK and CC α 1B calcium channels are analyzed along with the structural information, the footprint of SIGK upon G β is similar to the footprints of these signaling targets (Figure 4-12) (Ford, Skiba et al. 1998; Li, Sternweis et al. 1998; Agler, Evans et al. 2003). Of the thirteen residues from G β_1 that comprise the SIGK contact surface, nine (Lys57, Tyr59, Trp99, Met101, Leu117, Tyr145, Met188, Asp246, and Trp332) are also found as contacting residues in the G α , GRK2, and phosphoinositide-dependent kinase-1 complexes (Figure 4-13). These residues may be considered as a consensus set of residues utilized by many G β binding partners. An additional three of the thirteen residues (Asp186, Asp228, and Asn230) are shared among SIGK and two of the other protein complex structures. One of the thirteen, Val100, contacts SIGK through its main chain oxygen and is not involved in binding interactions in the other complexes. The SIGK binding residues that are most sensitive to mutational perturbation are also the most frequently involved in interactions with other G β binding partners (Figures 4-7 and 4-13). In addition, all of the residues in the SIGK binding interface are invariant across mammalian G β_1 -G β_4 isoforms (Figure 1-4). As discussed in Chapter One, G β isoforms are so highly conserved across species that the sequence conservation of these residues is not as relevant as their structural conservation (Figure 1-4 and Figure 4-13). SIGK was identified from a random peptide phage display where

multiple peptides unrelated by sequence appeared to bind to a common surface on $G\beta_1$. As shown in the $G\beta_1\gamma_2$ •SIGK structure this screen, unbiased in terms of biological function, selected a common $G\beta$ surface shared by multiple $G\beta$ signaling targets.

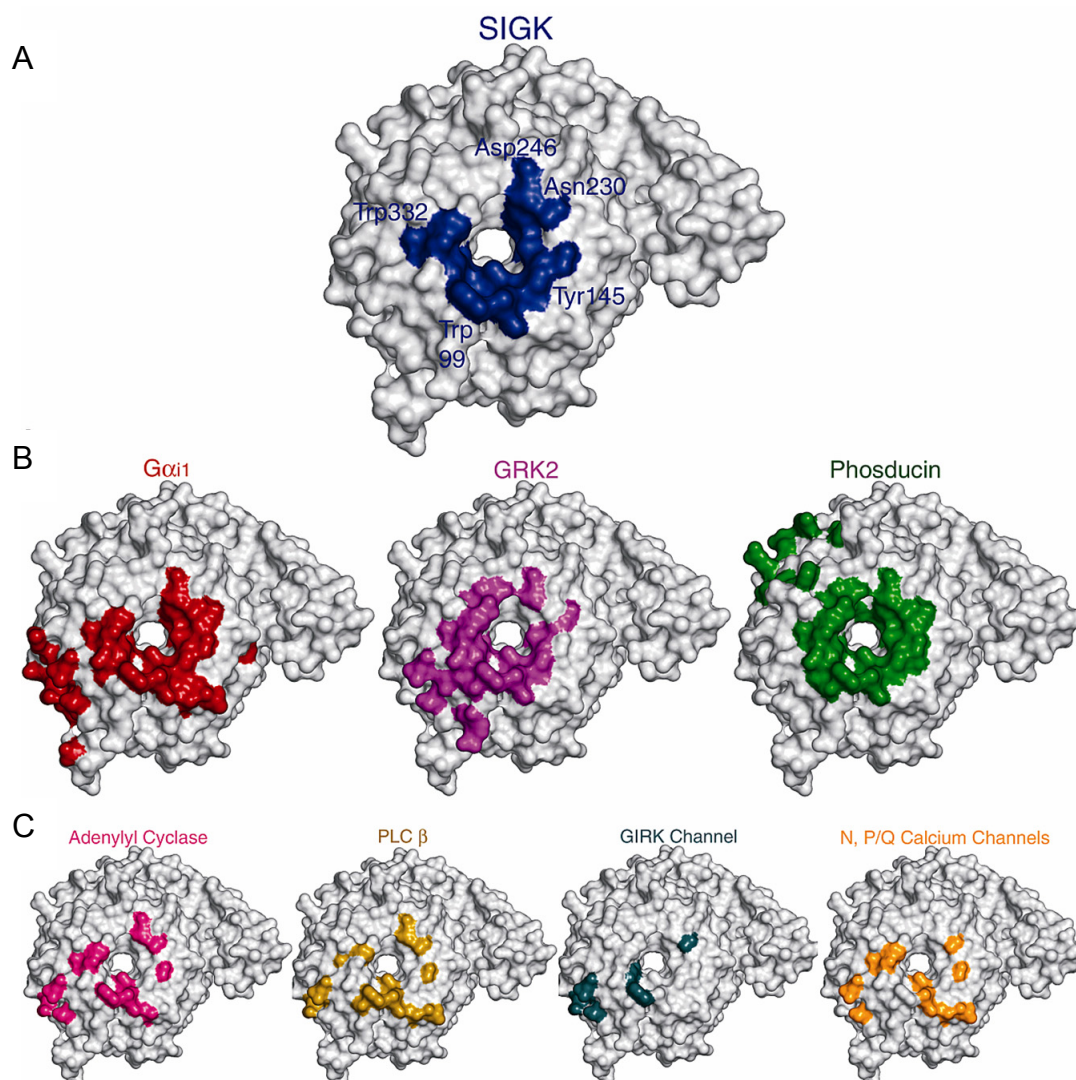


Figure 4-12. Molecular surface comparisons of $G\beta\gamma$ binding interactions.

Molecular surfaces are shown for $G\beta\gamma$; the residues in each color correspond to the $G\beta$ binding surface for each signaling target. (A) Residues that contact the SIGK peptide. Relative positions of key residues of the $G\beta_{1\gamma 2}$ •SIGK interaction discussed in the text are labeled on the surface for orientation purposes. (B) Residues indicated by complex crystal structures to be important for target binding (Gaudet, Bohm et al. 1996; Lambright, Sondek et al. 1996; Loew, Ho et al. 1998; Wall, Posner et al. 1998; Lodowski, Pitcher et al. 2003). (C) Residues proposed to be important for target binding to $G\beta\gamma$ based on mutagenesis studies (Ford, Skiba et al. 1998; Li, Sternweis et al. 1998; Agler, Evans et al. 2003). The $G\beta\gamma$ binding partner for each molecule is labeled.

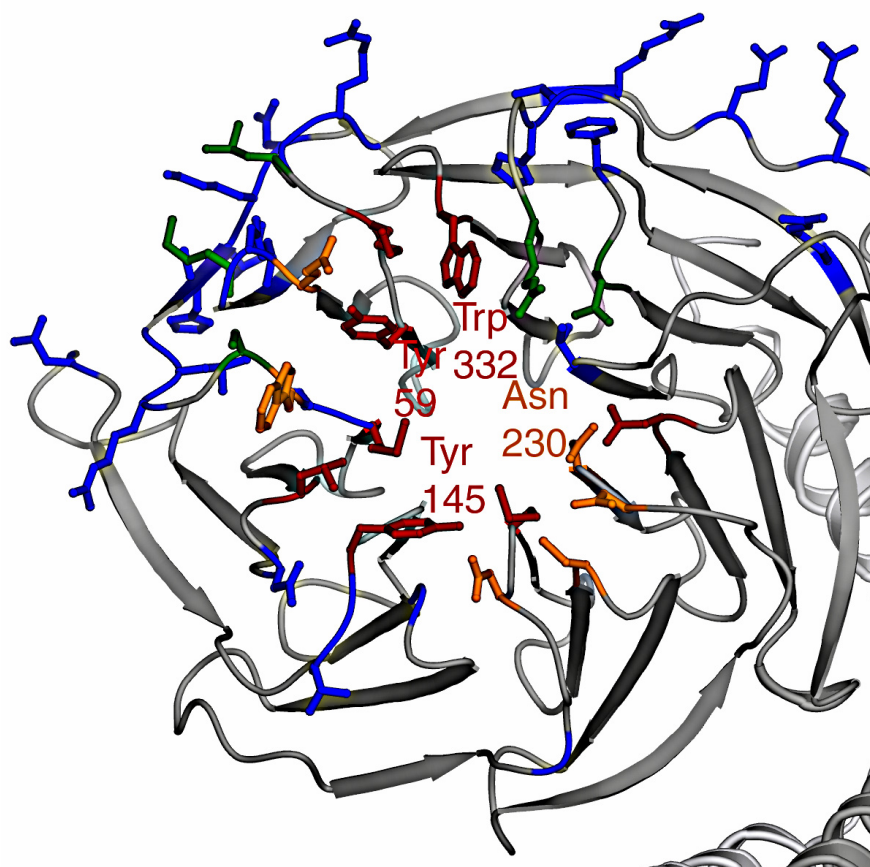


Figure 4-13. Structural conservation of G β at the SIGK binding site.

(A) Commonality of residues in G β_1 binding interfaces in complexes with SIGK, G α_{i1} , GRK2, and phosducin. Residues that contribute to one of the complex surfaces are colored blue; two complexes, green; three complexes, orange, and if found in all four surfaces, red. Four residues from the SIGK surface are labeled to define the orientation.

SIGK Effects on G $\beta\gamma$ Signaling Pathways

In view of the extensive overlap between the residues of G β_1 that are accessed by SIGK and those involved in the binding of protein G $\beta\gamma$ targets, SIGK should be a competitive inhibitor of multiple G $\beta\gamma$ binding reactions (Figure 4-12). Indeed, the

closely related SIRQ peptide has effects on several $G\beta\gamma$ dependent pathways; it blocks $G\beta\gamma$ mediated inhibition of PLC $\beta 2$ and PI3K in purified protein preparations, and a cell permeable version of SIRQ (myristoylated SIRQ) induces ERK I/II activation in intact cells (Scott, Huang et al. 2001). These effects are sensitive to mutations of residues in SIGK that interact with the surface of $G\beta$: sLys4, sAla5, sPhe6, sIle8, sLeu9, and sGly10 have all been shown by alanine scanning to be important for inhibition of PLC $\beta 2$ activation by $G\beta_{1\gamma 2}$ (Scott, Huang et al. 2001). In addition, sLeu9 is important for the ability of SIGK to activate MAPK pathways in cell culture (Goubaeva, Ghosh et al. 2003). However, SIRQ does not block inhibition of adenylyl cyclase type I or N-type Ca^{++} channel regulation, even though their footprints are quite similar to those of $G\beta$ and PLC $\beta 2$ (Figure 4-12) (Scott, Huang et al. 2001). Conversely, mutations in $G\beta$ that abrogate SIGK binding do not equally affect interaction with other $G\beta\gamma$ binding partners. For example, mutation of Leu117 to alanine decreases the ability of $G\beta\gamma$ to activate adenylyl cyclase type II and PLC $\beta 3$ and to bind GRK2 and SIGK, but has no effect on GIRK1/GIRK4 potassium channel activation, $CC\alpha 1B$ calcium channel activation, or PLC $\beta 2$ activation (Table 4-2) (Ford, Skiba et al. 1998). Similarly, mutation of Trp332 to alanine reduces affinity of $G\beta_{1\gamma 2}$ for SIGK and impairs stimulatory activity towards adenylyl cyclase type II, $CC\alpha 1B$ and both PLC $\beta 2$ and PLC $\beta 3$, but does not affect interaction with GRK2, activation of GIRK1/GIRK4, or inhibition of adenylyl cyclase type I (Ford, Skiba et al. 1998). Both Leu117 and Trp332 form part of the $G\alpha_t$ and $G\alpha_{i1}$ binding sites of $G\beta_1$

(Wall, Coleman et al. 1995; Lambright, Sondek et al. 1996; Wall, Posner et al. 1998)

and mutation of Leu117 also affects $G_{\alpha_{i1}}$ association with $G\beta_1\gamma_2$ (Ford, Skiba et al. 1998).

Gα_{i1}	Phosducin	GRK2	SIGK	PLCβ	AC	GIRK	Ca⁺⁺
52	42						
53	44						
55	46						
57	47						
59		55			55	55	55
75	57	57	Lys 57		57		
	59	59	Tyr 59			59	
	75	75					
		76					
78		78			78	78	78
80				80		80	80
88							
89				89	89	89	
90							
91							
92							
		95					
		96					
	98	98					
99	99	99	Trp 99	99	99	99	
			Val 100				
101	101	101	Met 101	101	101		101
117	117	117	Leu 117	117	117		
119				119	119		119
132							
143				143			143
144							
145	145	145	Tyr 145				
	162						
182							
186	186		Asp 186	186	186		186
188	188	188	Met 188				
204	204	204					
228	228		Asn 228	228	228	228	
230	230						
246	246	246	Asp 246	246			
	274						
	290	290					
	292						
	304						
	310						
	314	314					
332	332	332	Trp 332	332	332		332
41%	44%	44%	-----	54%	64%	43%	38%

Table 4-2. Interaction surfaces for G β_1 binding partners. Key to column headings: G α_{i1} , the crystal structure of the G α_{i1} •G $\beta_1\gamma_2$ heterotrimer (Wall, Coleman et al. 1995; Wall, Posner et al. 1998); phosducin, the phosducin•G $\beta_1\gamma_1$ complex

(Gaudet, Bohm et al. 1996); GRK2, the GRK2•G $\beta_1\gamma_2$ complex (Lodowski, Pitcher et al. 2003); SIGK, the SIGK•G $\beta_1\gamma_2$ complex; PLC β , mutational analysis of the PLC $\beta_2/3$ •G $\beta_1\gamma_2$ complexes (Ford, Skiba et al. 1998; Li, Sternweis et al. 1998); AC, mutational analysis of the adenylyl cyclase type I/II•G $\beta_1\gamma_2$ complex (Ford, Skiba et al. 1998); GIRK, mutational analysis of G $\beta_1\gamma_2$ interaction with the GIRK1/4 channels (Ford, Skiba et al. 1998); Ca⁺⁺, mutational analysis of G $\beta_1\gamma_2$ interaction with N or P/Q type calcium channels (Ford, Skiba et al. 1998; Agler, Evans et al. 2003). In **bold** are residues which are important for the SIGK•G $\beta_1\gamma_2$ interaction. The last row calculates the percentage of residues that are shared between the target and SIGK interfaces.

The failure of SIGK to act as a universal G β competitor must be attributed to the importance of residues that lie outside of the “consensus” site for stable binding of some G β binding partners. The contribution of these extra-consensus sites to the binding energy may differ among G β -interacting proteins. G α_i , for example, interacts with G $\beta\gamma$ at two distinct loci. The switch II region binds to the consensus site, and the N terminal helix of G α_{i1} binds along the side of the G β_1 torus (Wall, Coleman et al. 1995; Lambright, Sondek et al. 1996; Wall, Posner et al. 1998). Both interfaces contribute to the energy of binding; N terminally truncated G α_{i1} fails to bind G $\beta_1\gamma_2$, and the two interfaces bury similar amounts of solvent accessible surface area (Wall, Coleman et al. 1995). Thus, it is possible that SIGK could bind to G protein heterotrimers upon dissociation of the switch II region of G α_i from G β , even while the N terminus of G α_i remains associated with G $\beta\gamma$. Stimulation of heterotrimer dissociation by SIKR may involve formation of such an intermediate complex (G α_{i1} •G $\beta\gamma$ •SIGK). The G $\beta_1\gamma_2$ •SIGK structure shows that SIGK does not induce substantial conformational change in G β_1 outside of the SIGK binding site itself

(Figure 4-8). Therefore, based solely upon the $G\beta_{1\gamma 2}$ •SIGK complex structure there is no explanation for the molecular mechanism of SIGK induced $G\alpha$ dissociation.

Other $G\beta$ binding partners might form strong interactions with extra-consensus residues such that binding of competitors at the consensus site would not be sufficient to affect $G\beta$ binding. For example, peptides which bind the top face of $G\beta$, including the QEHA peptide derived from adenylyl cyclase type II and SIRK/SIGK peptide, are able to compete directly for binding of $G\alpha$ to $G\beta$. This indicates that at least some adenylyl cyclase isoforms share an interface with SIGK on the top face of $G\beta$ (Ghosh, Peterson et al. 2003). However, mutagenesis experiments did not provide evidence for binding of adenylyl cyclase type I to the consensus surface of $G\beta_1$, which is consistent with the inability of SIRK to interfere with $G\beta\gamma$ inhibition of this effector (Li, Sternweis et al. 1998). Perhaps this discrepancy may be explained by the unique extra-consensus surface shared by adenylyl cyclase isoforms and $G\beta\gamma$. It should be noted that effect of SIGK on the formation of other $G\beta_1$ complexes, for instance those with GRK2 and phosducin, has not been tested.

Analysis of the Top Face of $G\beta$ as a “Hot Spot”

The ability of the SIGK binding site of $G\beta_{1\gamma 2}$ to recognize a range of protein ligands with diverse secondary structures (Figures 4-9 through 4-12) suggests that it may be an example of a preferential protein binding site as described by Delano et al. (DeLano, Ultsch et al. 2000). Preferential binding surfaces are characterized as

having high solvent accessibility, low polarity, and a large degree of conformational flexibility (Clackson and Wells 1995; Bogan and Thorn 1998; Ma, Wolfson et al. 2001; Scott, Huang et al. 2001; DeLano 2002). Moreover, preferential binding sites are likely to contain an unusually high concentration of so-called “hot spots”: residues that, if mutated to alanine, reduce binding energy at least ten-fold (DeLano 2002). Hot spots have been described for both protein-protein and protein-small molecule interfaces; often point mutations to any hot spot on a surface completely abrogate complex formation, even when the binding interfaces bury several hundred Å² of total surface area (Clackson and Wells 1995; Bogan and Thorn 1998; Thanos, Randal et al. 2003; Zhang, Li et al. 2003). We have used these criteria to evaluate the SIGK binding site of Gβ as a protein surface that is predisposed by its chemical composition and surface properties to serve as a protein binding site.

Of the twelve residues in the SIGK contact surface that were tested in the current study, eight (Lys57, Tyr59, Leu117, Tyr145, Asp186, Met188, Asn230, and Trp332) met the energetic criterion for a “hot spot” residue. Replacement of any of these residues by alanine resulted in a 10-fold reduction in the affinity of Gβ_{1γ2} for SIGK (Figure 4-5 and Figure 4-7). It is clear that all of these residues act as energetically important nodes that contribute favorably to SIGK binding.

The SIGK binding surface of Gβ contains several residues that Bogan and Thorn have shown to be enriched in hot spots (Bogan and Thorn 1998). These include

tyrosine, tryptophan and arginine; bulky residues that are capable of forming both polar and non-polar interactions. This preference for aromatic and extended alkyl chain residues leads to the classification of hot spots as nonpolar cores surrounded by more solvent-accessible residues. The SIGK contact surface contains residues that are underrepresented in hot spot regions such as Leu117 and Val100, along with Met101 and Met188. However, the SIGK binding surface is significantly more populated with aromatic residues (Figure 4-4 and Figure 4-6) than the rest of the G β surface. 38% of the SIGK binding surface versus 8.5% of the total non-glycine surface accessible G β residues is composed of Phe, Tyr, His, or Trp. Therefore, the SIGK surface is more nonpolar; in total, 62% of the SIGK binding surface is nonpolar compared to 29% of G β surface accessible residues. It is interesting to note that asparagine and aspartic acid, which have a moderately favorable distribution among hot spot surfaces, account for four of the thirteen residues in the SIGK binding surface. This combination of aromatic and charged residues allows for accommodation of binding partners with diverse chemical properties at the G β binding surface, although it does not fit the literature definition of a hot spot mode of interaction.

Preferential binding surfaces are expected to have high surface accessibility (DeLano, Ultsch et al. 2000). To analyze this property of the SIGK binding surface, the total surface accessible area was calculated for the G β molecule on a residue, main chain, and side chain basis. Accessibility scores for SIGK-binding residues

relative to residues of the same type in $G\beta_1$ are shown in Figure 4-14. Most amino acids in the SIGK binding surface are not significantly more accessible than others of their type in $G\beta$. However, five residues showed significant deviation from the mean: Tyr59, Trp99, Met101, Leu117, and Trp332. In the case of Trp99 side chain surface accessibility is significantly greater than the type average; the main chain of Tyr59, Trp99, Met101, and Trp332 are more accessible than the mean. Leu117 has significantly large main chain and side chain accessibility.

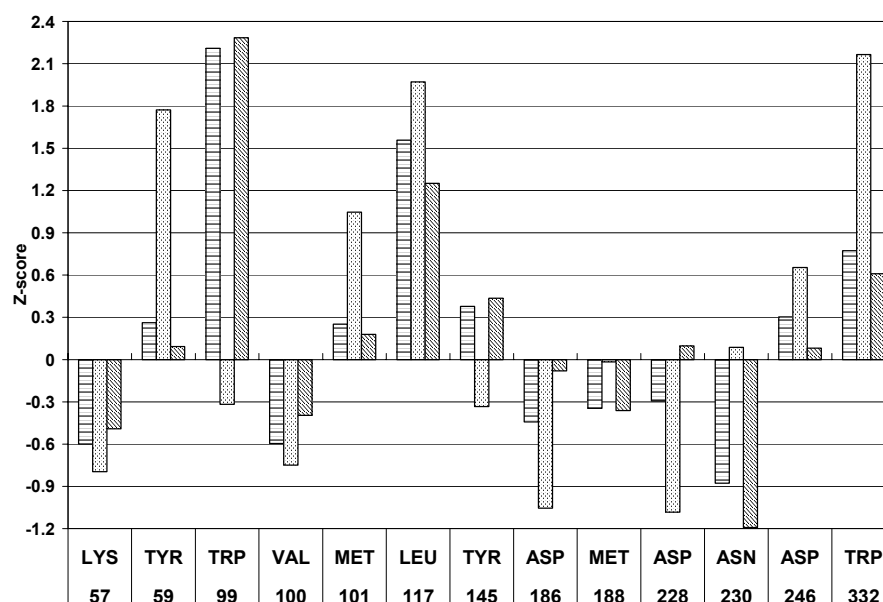


Figure 4-14. Solvent accessibility of the SIGK binding site on $G\beta_1$. Total residue, main chain, and side chain solvent accessible area (\AA^2) were computed for all $G\beta_1$ residues in the $G\beta_{1\gamma_2}$ •SIGK, $G\beta_{1\gamma_2}$ • $G\alpha_{i1}$, $G\beta_{1\gamma_2}$ •GRK2, and $G\beta_{1\gamma_1}$ •phosducin complexes using CNS (Lee and Richards 1971). Average and standard deviations were calculated for each of the twenty amino acid types. A solvent accessibility Z-score was computed for each residue, n , at the SIGK interface: $Z_n = (A_n - A_i) / \sigma_i$, where A_n is the solvent accessibility of residue n , i represents the side chain type of that residue, and A_i and σ_i are the mean and standard deviation of solvent accessibility for residues of type i . Bars filled with horizontal stripes represent accessibility data for all atoms of the residue; dotted bar, main chain atoms only; diagonal stripe, side chain atoms.

Conformational flexibility or adaptability has been cited as an important determinant of a preferential binding surface, since such surfaces are better able to bind to structurally unrelated protein targets (DeLano, Ultsch et al. 2000). Residue flexibility can be quantified in terms of relative positional variation in the context of several protein complexes as shown in Figure 4-15. A histogram showing the RMSD relative to uncomplexed $G\beta_1\gamma_1$ of all $G\beta$ residues in four crystal structures ($G\beta_1\gamma_2$ •SIGK; $G\beta_1\gamma_2$ • $G\alpha_{i1}$; $G\beta_1\gamma_2$ •GRK2; $G\beta_1\gamma_1$ •phosducin) shows that the SIGK surface residues of $G\beta_1$ show only slightly greater than average side chain positional dispersity (1.42 Å compared to 1.35 Å), with the side chains of Trp99, Asp228, and Trp332 having the largest positive deviation from the average (each greater than 2 Å). In particular, Arg314 and Trp332 in blade 7 move more than 10 Å towards the outside of the $G\beta_1$ torus to interact with phosducin.

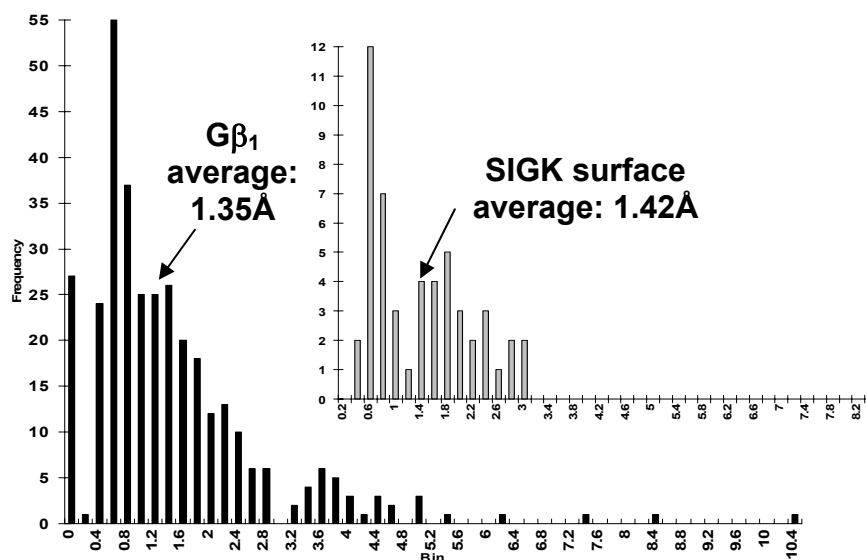


Figure 4-15. Positional variation of $G\beta_1$ residues in the SIGK binding surface. The RMSD (in Å) of $G\beta_1$ residues in $G\alpha_1$ -protein complexes with respect to their positions in unbound $G\beta_1\gamma_1$ (after superposition as shown in Figures 4-9 through 4-11) are plotted as a histogram. The mean RMSD for all $G\beta_1$ residues and the mean RMSD for the thirteen residues in the SIGK binding surface are marked with arrows. The histogram for the thirteen SIGK binding residues is shown as an inset.

Atomic B factors also provide a measure of conformational flexibility. In the structure of uncomplexed $G\beta_1\gamma_1$ the B factors for Trp99, Val100, and Met101 exceed the mean value by least one standard deviation (Trp99 is greater than 2 standard deviations from the mean). However, in complexes with $G\alpha_{i1}$, GRK2, phosducin, and SIGK complexes, these binding site residues become more well ordered with B values close to the mean and in some cases up to 1 standard deviation below the mean. Thus, the capacity of $G\beta$ to recognize structurally diverse binding partners does not require a high degree of conformational flexibility for most residues in the top binding surface. Small structural adaptations are sufficient to accommodate a range of co-evolving binding partners in the $G\beta_1$ molecule.

Structural and mutagenic analysis demonstrates that the SIGK binding site on $G\beta_1$ may be regarded as a hot surface, co-evolved to promote tight binding with multiple protein targets. However, the mechanism by which $G\beta\gamma$ acts as a hot surface is complex. Trp332 is the only residue which meets all four of the criteria for a “hot spot”, although Tyr59 and Trp99 have three of the four characteristics of hot spot residues that were tested. There are other residues in the top face of $G\beta$ that are sensitive to mutational perturbation and are utilized in many binding partner interactions, but do not show significant shifts in conformational flexibility, solvent accessibility, or nonpolarity. Especially notable among this group are Lys57 and Met188; both of these residues are clearly important in the $G\beta$ interface as shown by mutational analysis and comparison to known $G\beta\gamma$ complex structures, and yet do not meet any of the additional statistical criteria for hot spot residues. Like other proteins that are required to interact with multiple, structurally disparate binding partners (DeLano, Ultsch et al. 2000), $G\beta$ has evolved a highly sensitive and specialized surface for interaction. In this case, the molecular mechanism mainly depends on heterogeneous non-bonded interactions at the binding surface; residues are utilized which use van der Waals contacts (methionine and leucine), polar contacts (aspartate and asparagine), and both (lysines, tryptophan, and tyrosine) to contribute energy to the binding interface.

BIBLIOGRAPHY

- Agler, H. L., J. Evans, et al. (2003). "Custom distinctions in the interaction of G-protein beta subunits with N-type (CaV2.2) versus P/Q-type (CaV2.1) calcium channels." J Gen Physiol **121**(6): 495-510.
- Bogan, A. A. and K. S. Thorn (1998). "Anatomy of hot spots in protein interfaces." J Mol Biol **280**(1): 1-9.
- Clackson, T. and J. A. Wells (1995). "A hot spot of binding energy in a hormone-receptor interface." Science **267**(5196): 383-6.
- DeLano, W. L. (2002). "Unraveling hot spots in binding interfaces: progress and challenges." Curr Opin Struct Biol **12**(1): 14-20.
- DeLano, W. L., M. H. Ultsch, et al. (2000). "Convergent Solutions to Binding at a Protein-Protein Interface." Science **287**(5456): 1279-1283.
- Ford, C. E., N. P. Skiba, et al. (1998). "Molecular basis for interactions of G protein betagamma subunits with effectors." Science **280**(5367): 1271-4.
- Gaudet, R., A. Bohm, et al. (1996). "Crystal structure at 2.4 angstroms resolution of the complex of transducin betagamma and its regulator, phosducin." Cell **87**(3): 577-88.
- Ghosh, M., Y. K. Peterson, et al. (2003). "Receptor- and nucleotide exchange-independent mechanisms for promoting G protein subunit dissociation." J Biol Chem **278**(37): 34747-50.
- Goubaeva, F., M. Ghosh, et al. (2003). "Stimulation of cellular signaling and G protein subunit dissociation by G protein betagamma subunit-binding peptides." J Biol Chem **278**(22): 19634-41.
- Kozasa, T. and A. G. Gilman (1995). "Purification of recombinant G proteins from Sf9 cells by hexahistidine tagging of associated subunits. Characterization of alpha 12 and inhibition of adenylyl cyclase by alpha z." J Biol Chem **270**(4): 1734-41.
- Lambright, D. G., J. Sondek, et al. (1996). "The 2.0 A crystal structure of a heterotrimeric G protein." Nature **379**(6563): 311-9.

- Lee, B. and F. M. Richards (1971). "The interpretation of protein structures: estimation of static accessibility." J Mol Biol **55**(3): 379-400.
- Li, Y., P. M. Sternweis, et al. (1998). "Sites for Galpha binding on the G protein beta subunit overlap with sites for regulation of phospholipase Cbeta and adenylyl cyclase." J Biol Chem **273**(26): 16265-72.
- Lodowski, D. T., J. A. Pitcher, et al. (2003). "Keeping G proteins at bay: a complex between G protein-coupled receptor kinase 2 and Gbetagamma." Science **300**(5623): 1256-62.
- Loew, A., Y. K. Ho, et al. (1998). "Phosducin induces a structural change in transducin beta gamma." Structure **6**(8): 1007-19.
- Ma, B., H. J. Wolfson, et al. (2001). "Protein functional epitopes: hot spots, dynamics and combinatorial libraries." Curr Opin Struct Biol **11**(3): 364-9.
- Nicholls, A., K. A. Sharp, et al. (1991). "Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons." Proteins **11**(4): 281-96.
- Rice, P., I. Longden, et al. (2000). "EMBOSS: the European Molecular Biology Open Software Suite." Trends Genet **16**(6): 276-7.
- Sarvazyan, N. A., A. E. Remmers, et al. (1998). "Determinants of gi1alpha and beta gamma binding. Measuring high affinity interactions in a lipid environment using flow cytometry." J Biol Chem **273**(14): 7934-40.
- Scott, J. K., S. F. Huang, et al. (2001). "Evidence that a protein-protein interaction 'hot spot' on heterotrimeric G protein betagamma subunits is used for recognition of a subclass of effectors." Embo J **20**(4): 767-76.
- Sondek, J., A. Bohm, et al. (1996). "Crystal structure of a G-protein beta gamma dimer at 2.1A resolution." Nature **379**(6563): 369-74.
- Thanos, C. D., M. Randal, et al. (2003). "Potent small-molecule binding to a dynamic hot spot on IL-2." J Am Chem Soc **125**(50): 15280-1.
- Wall, M. A., D. E. Coleman, et al. (1995). "The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2." Cell **83**(6): 1047-58.
- Wall, M. A., B. A. Posner, et al. (1998). "Structural basis of activity and subunit recognition in G protein heterotrimers." Structure **6**(9): 1169-83.

Zhang, L., Z. Li, et al. (2003). "Mutagenesis of the Runt domain defines two energetic hot spots for heterodimerization with the core binding factor beta subunit." J Biol Chem **278**(35): 33097-104.

CHAPTER FIVE

Crystallization of $G_{\alpha_{i1}}$ Bound to Fluorophore Nucleotides: First Attempt

A project involving G_{α} began at the same time as the SIGK• $G_{\beta_1\gamma_2}$ project described in Chapters Two through Chapter Four. The goal of this project was to crystallize a complex between G_{α} and various fluorophore labeled nucleotides. This chapter provides an introduction to the goals of the project; the methods used for protein purification, crystallization, and structure determination; and the results obtained. The first attempt to purify untagged $G_{\alpha_{i1}}$ failed; instead, a contaminant protein was repeatedly purified and crystallized. In this chapter, the procedures used to determine the structure of a contaminant are outlined, and the structure of the contaminant is shown. At the end of the chapter, strategies to eliminate this contaminant from the $G_{\alpha_{i1}}$ purification and improvements to the exchange and crystallization procedures are discussed.

The use of Fluorophore Nucleotides to Elucidate G Protein Signaling

Fluorophore labeled nucleotides are a useful tool for enzymatic assays and for labeling specific proteins in cells or tissues. The sensitivity of fluorescence can be advantageous in detecting small amounts of proteins in cellular assays. For fluorophores with excitation wavelengths in the proper range, intrinsic fluorescence can be amplified by FRET within the protein, thereby leading to specific fluorescence only when the fluorophore is bound. In addition, when the fluorophore modification does not affect the binding properties or kinetic parameters of the substrate,

fluorescent monitoring of enzymatic processes is often preferred by researchers to radioactive substrates. Finally, fluorescence can be used to track particular processes in cells and tissues, and multiple fluorophores can be utilized to simultaneously observe separate events in a larger system.

MANT (2'(3')-O-(*N*-methylantraniloyl) substituted nucleotides (Figure 5-1) have been utilized to study small G proteins, and the binding affinities and rate constants for MANT-GTP and MANT-GDP are comparable to unlabeled nucleotide for p21Ras (Remmers, Posner et al. 1994). More recently, MANT labeled nucleotides have been used in multiple studies of heterotrimeric G proteins, with varied results (Remmers, Posner et al. 1994; Remmers and Neubig 1996; Remmers 1998; Gille and Seifert 2003). The affinities for unlabeled nucleotides can be represented with the rank order $\text{GTP}\gamma\text{S} > \text{GMMPNP} = \text{GTP} = \text{GDP} > \text{GMP}$; for MANT, the order is $\text{MANT-GTP}\gamma\text{S} > \text{MANT-GMMPNP} > \text{MANT-GTP}$ (Bokoch, Katada et al. 1984; Remmers and Neubig 1996). There are differences in binding affinities between isoforms of $\text{G}\alpha$; most notably, $\text{G}\alpha_t$ does not bind to MANT-GTP γ S. The affinity of myristolated $\text{G}\alpha_i$ for MANT-GTP γ S was originally reported to be nearly identical to unlabeled GTP γ S ($\text{EC}_{50} = 11 \text{ nM}$ vs. 22 nM) (Remmers and Neubig 1996). However, later studies with receptor-linked $\text{G}\alpha_i$ isoforms indicate a higher K_i for MANT-GTP γ S compared to unlabeled GTP γ S (10 nM vs. 1090 nM) (Gille and Seifert 2003; Gille and Seifert 2003). The reported K_d of MANT labeled GMMPNP for myristolated $\text{G}\alpha_i$ is 1500 nM ; affinities are reported to be lower for unmyristolated

$G\alpha_i$ but have not been quantified in the literature to date (Remmers and Neubig 1996). Off rates of MANT-GDP from $G\alpha_i$ have not been reported, but for $G\alpha_o$ dissociation of MANT-GDP is faster by two orders of magnitude compared to unlabeled GDP ($t_{1/2}=1.7s$ compared to 120s) (Remmers, Posner et al. 1994). MANT compounds have been used successfully to label G proteins in receptor-linked systems in insect cells, albeit with low affinity and some cross-reactivity towards adenylyl cyclase (Gille and Seifert 2003; Gille and Seifert 2003), and have been utilized in various enzymatic studies with partially or wholly purified proteins (Remmers, Posner et al. 1994; Remmers and Neubig 1996; Remmers 1998; Remmers, Engel et al. 1999).

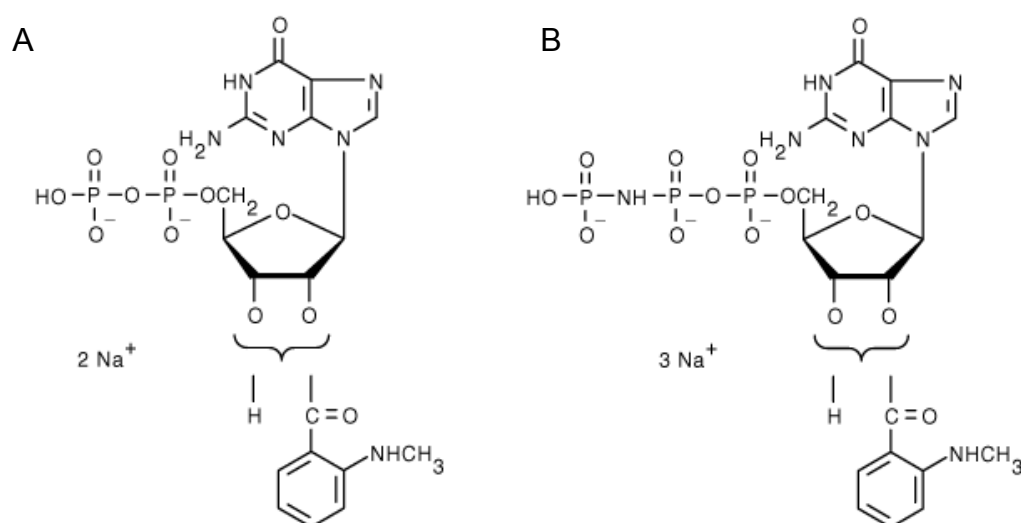


Figure 5-1. MANT labeled nucleotides used in the current study. (A) MANT-GDP. (B) MANT-GMPPNP. The MANT moiety is coupled nonspecifically to the 2' or 3' hydroxyl of the ribose ring of the nucleotide; exchange is fast and reversible between the two isoforms, although in most enzymatic assays only one form of the nucleotide has high affinity.

A different type of fluorophore, unrelated in structure, has also been linked to guanine nucleotides (Figure 5-2). BODIPY (4,4-difluoro-4-bora-3a,4a-diaza-s-indacene) has the advantage over MANT of being excited in the visible range of the spectrum, thereby lessening the effects of heating and background fluorescence seen in systems illuminated in the UV range. In addition, BODIPY has a large fluorescence enhancement upon binding $G\alpha$ that does not rely on FRET, as MANT fluorescence does (McEwen, Gee et al. 2001). Instead, the BODIPY moiety most likely becomes exposed to the environment upon $G\alpha$ binding, leading to a release of quenching by guanine (McEwen, Gee et al. 2001). The K_d of BODIPY-GTP γ S for $G\alpha_i$ has been reported as 150 nM; for BODIPY-GMPPNP the K_d is 10-fold lower (1500 nM) (McEwen, Gee et al. 2001). However, in a receptor coupled $G\alpha$ system, the K_i of BODIPY-GTP γ S and BODIPY-GMPPNP was reported to be even lower than for MANT nucleotides (11 μ M and 8.7 μ M, respectively) (Gille and Seifert 2003). The same study found that although the affinities of BODIPY nucleotides are greatly reduced compared to unlabeled or MANT labeled nucleotides, BODIPY-GTP γ S and BODIPY-GMPPNP were able to activate $G\alpha_i$ as efficiently as unlabeled GTP γ S and GMPPNP in an adenylyl cyclase activity assay (Gille and Seifert 2003). Recently, BODIPY labeled GTP γ S was utilized in an affinity probe capillary electrophoresis (APCE) assay to analyze mixtures of $G\alpha_o$ and $G\alpha_i$, or mixtures of Ras and Rab3A (Jameson, Cunliffe et al. 2003). These experiments are useful to

determine reaction kinetics, or to quantify classes of proteins based on their affinity to ligand from a complex mixture of analytes (Jameson, Cunliffe et al. 2003).

Current structures of $G\alpha$ bound to GDP and GTP analogs do not explain the differences in reaction kinetics seen in the MANT and BODIPY labeled analogs. In addition, the relatively high binding of BODIPY-GTP γ S is surprising, considering the large fluorescent group protruding from the γ thiol, and this cannot be explained using current knowledge. Finally, the structures of $G\alpha$ protein bound to these analogs might provide structural clues that could lead to the design of fluorescent nucleotides with more desirable kinetic parameters for multiple $G\alpha$ isoforms.

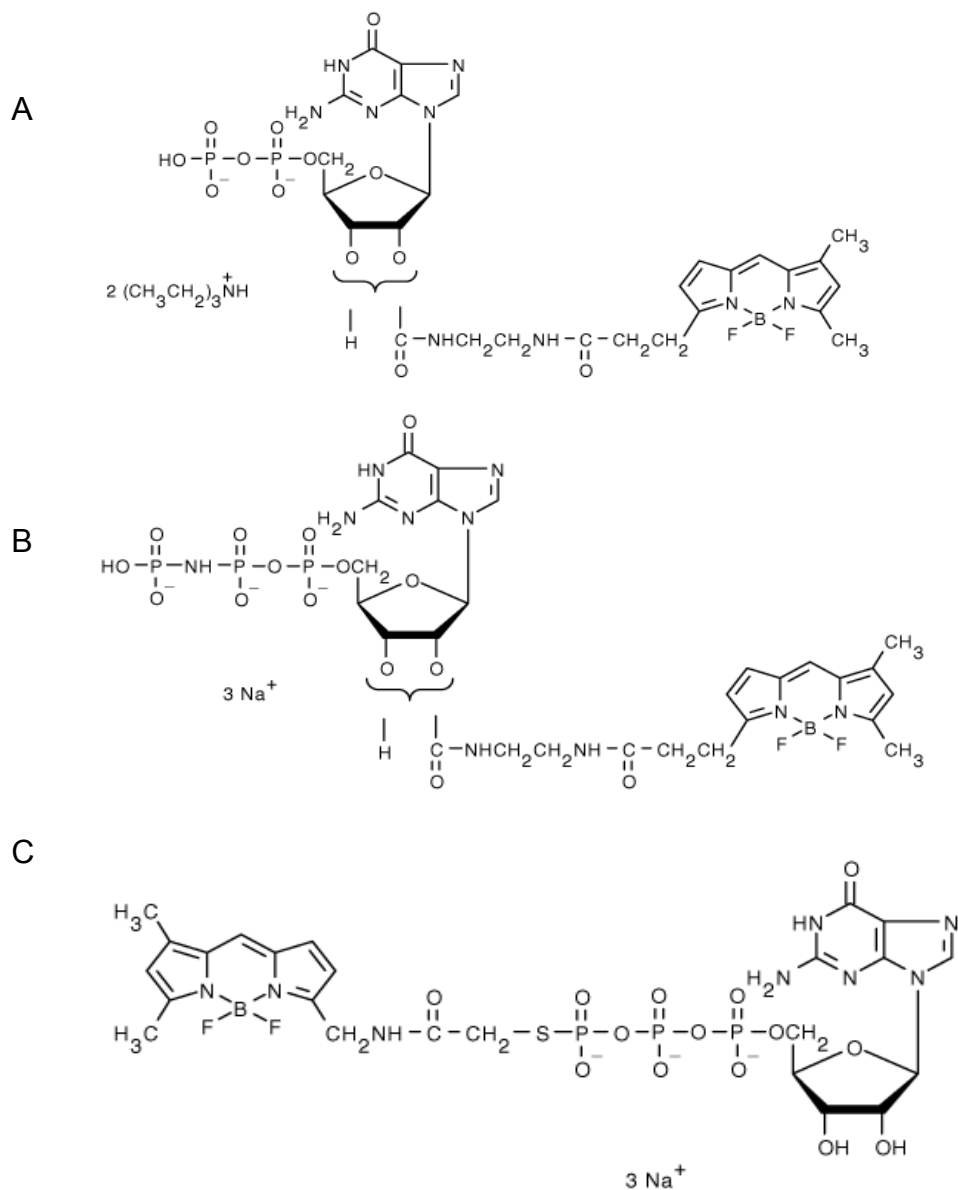


Figure 5-2. BODIPY labeled nucleotides used in the current study. (A) BODIPY-GDP. (B) BODIPY-GMPPNP. (C) BODIPY-GTP γ S. BODIPY is coupled to GDP and GMPPNP through the ribose ring, as in the MANT nucleotides. The GTP γ S nucleotide is coupled to BODIPY through the γ thiol.

MATERIALS AND METHODS

Expression and Purification of Untagged $G\alpha_{i1}$

Untagged PQE-60 plasmid harboring $G\alpha_{i1}$ was used to drive expression of untagged protein. The construct is identical to that in (Lee, Linder et al. 1994). Expression was carried out using *e. coli* strain JM109. Cells were grown to an $OD_{595}=0.4$ at 30 °C; 30 μ M IPTG was added to cells and cells were allowed to grow for 12-16 hours at 30 °C. Cells were harvested at 2600g for 30 minutes and lysed by sonication (5 s on, 5 s off, 5 minutes total sonication time). Protease inhibitors (1 mL per 6L of cell culture of Sigma P-2714) and DNase I (5 mg +2 mM $MgCl_2$) was added to the lysed cells, and the supernatant was clarified by ultracentrifugation at 125,000g for 45 minutes. Supernatant was diluted to 300 mLs using buffer QA (20 mM Tris pH 8, 5 mM β -mercaptoethanol), loaded onto a 50 mL Q superose column and eluted with 10 column volumes of buffer QB(QA + 0.5M NaCl). The fractions were tested by SDS-PAGE; fractions containing $G\alpha_{i1}$ were pooled and dialyzed against buffer HAP-A (10 mM Tris pH 7.2, 5 mM β -ME, 0.01M NaCl) overnight. The next day, the eluate was run over a 20 mL hydroxyapatite column and eluted off in 10 column volumes of buffer HAP-B (10 mM Tris pH 7, 5 mM β -ME, 0.01M NaCl, and 0.25M KPO_4 pH 7). The fractions were tested using SDS-PAGE, and fractions containing $G\alpha_{i1}$ were pooled, brought to a final concentration of 1.2 M ammonium sulfate, and then filtered through a 0.2 μ filter. This eluate was then loaded on a 10 mL phenyl superose column pre-equilibrated in buffer PHE-A (50 mM Tris pH 8, 3 mM EDTA, 5 mM β -

ME, and 1.2 M $(\text{NH}_4)\text{SO}_4$), and eluted off in 10 column volumes of PHE-B (50 mM Tris pH 8, 3 mM EDTA, and 5 mM β -ME). The fractions were analyzed using SDS-PAGE for $\text{G}\alpha_{i1}$, and the eluate pooled and concentrated. The protein buffer was exchanged on a gel filtration column (Sephadex 200:75) into buffer GF (20 mM HEPES pH 8, 150 mM NaCl, 10 mM β -ME, and 1mM EDTA); fractions were analyzed using SDS-PAGE and the peak fractions collected and concentrated for further experiments. Typical yields were 10 mg protein/L cell culture.

Exchange of Fluorophore Nucleotide for Endogenous GDP

Exchange was accomplished by adding 100 mM ammonium sulfate, 2 mM MgCl_2 , and 5x molar excess of the nucleotide to concentrated $\text{G}\alpha_i$. Typical concentrations of protein used in the experiment ranged from 25-250 μM protein. The resulting solution was placed at 30 °C for three hours. The protein mixture was then diluted with at least 10 volumes of buffer GF and concentrated to at least 10 mg/mL.

Crystallography of Untagged $\text{G}\alpha_{i1}$ •Fluorophore Complex

After concentration, 2 mM MgCl_2 and 1.5 molar excess fluorescent nucleotide was added to the protein solution. Crystal conditions were based on those used previously. (2-2.5 M ammonium sulfite, 100 mM sodium acetate pH 6.0) (Coleman, Lee et al. 1994).

Molecular Replacement

Molecular replacement trials were carried out using the programs CNS and AMoRE as outlined in Chapter Three. Multiple search models were used, including those of $G\alpha_i$ complexed with GDP (1GDD) (Mixon, Lee et al. 1995), GMPPNP (1CIP) (Coleman and Sprang 1999), GTP γ S (1GIA) (Coleman, Berghuis et al. 1994), and aluminum fluoride (1GFI) (Coleman, Berghuis et al. 1994). Search models where loops and protein termini were removed were also used.

Expression and Purification of Selenomethionine Labeled Untagged $G\alpha_{i1}$

Selenomethionine labeled $G\alpha_{i1}$ was expressed using the same plasmid carrying the $G\alpha_{i1}$ gene as outlined above. However, an auxotrophic cell line (BL21 (DE₃)) was utilized to produce selenomethionine labeled protein. Induction conditions were identical to that of unlabeled protein, although the media formulation used to grow the cells was different and is shown in Table 5-1. Purification of selenomethionine labeled $G\alpha_i$ was identical to purification of unlabeled protein. Mass spectrometry was used to calculate percent selenomethionine incorporation into the protein. Typical yields were 5 mg protein/L cell culture. Nucleotide exchange conditions were identical for selenomethionine labeled and unlabeled $G\alpha_{i1}$.

Autoclaveable Ingredients	
Ingredient	Grams per L
Ammonium Sulfate	1
KH ₂ PO ₄	4.5
K ₂ HPO ₄	10.5
Na-Citrate•2H ₂ O	0.5
NaCl	0.5
Amino acids:	
Ala	0.8
Arg	0.8
Asp	0.8
Asn	0.8
Cys	0.8
Gln	0.8
Glu	0.8
Gly	0.8
His	0.8
Ile	0.8
Leu	0.8
Lys	0.8
Phe	0.8
Pro	0.8
Ser	2.2
Thr	0.8
Trp	0.8
Tyr	0.8
Val	0.8
Nucleosides:	
Adenine	0.6
Guanosine	0.7
Thymine	0.125
Uracil	0.6
Non-Autoclaveable Ingredients	
Ingredient	Grams per L
D-Glucose	10
MgSO ₄	0.25 (1 mL of 1M)
FeSO ₄	0.0835
Thiamine	0.04
Biotin	2 mL of 2 mg/mL (tris8)
Se-met	0.05

Table 5-1. Media formulation used in selenomethionine experiments. The selenomethionine is added at the same time the cells are added to the media, and may be supplemented at the time of protein induction to ensure full incorporation.

Crystallography of Selenomethionine Labeled Untagged $G\alpha_{i1}$

Crystallography of selenomethionine labeled $G\alpha_i$ was identical to that for unlabeled $G\alpha_{i1}$. Crystals were slightly smaller and diffracted to lower resolution than unlabeled protein. Fluorescence scans were used at beam line ID-19 at APS to verify selenomethionine incorporation in the crystals; all datasets were then collected on optimized crystals at 8.2.1 and 8.2.2 at ALS.

MAD Data Collection and Phase Calculation

As discussed in Chapter Three, there are several ways to solve the phase problem inherent in x-ray diffraction studies. MAD phasing takes advantage of the anomalous dispersion of x-rays seen for certain metals with resonances in the range of wavelengths seen at synchrotron sources (Hendrickson 1991). Selenium is one of these metals, with absorption values in the range of 0.95-1Å (12-13 keV) (Hendrickson 1991). Selenium can be incorporated into the protein of interest by inducing the protein under minimal media conditions containing selenium-substituted methionine. In many cases, the protein expression, purification, and crystallization protocols need not be modified after selenomethionine incorporation (Doublié 1997).

During a typical x-ray experiment, atomic scattering factors (f_j) include a component that is contributed by the electron density distribution of the atom, and a dispersive component that is related to the bound electronic state of the atom. In a MAD

experiment, the wavelength of the incident x-ray is chosen such that it creates resonant absorption of energy in the anomalous center (Hendrickson 1991) (Figure 5-3). This experiment leads to the collection of anomalous scattering component (f') to the atomic scattering factor.

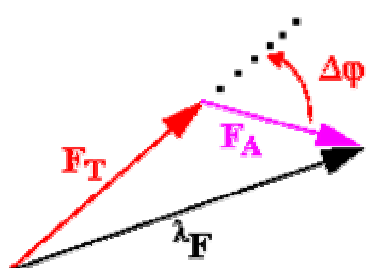


Figure 5-3. Phase diagram for MAD experiments.

Key to figure: λF – total scattering amplitude at a given wavelength; phase is unknown. F_T – normal scattering of all atoms; F_A – anomalous scattering component of all atoms; $\Delta\phi$ = difference in phase angle between normal and anomalous scattering components. Taken from

http://www.bmsc.washington.edu/scatter/MAD_1.html .

Since there is a contrast between the scattering from anomalous centers compared to the rest of the atoms in the protein, the positions of the anomalous scattering atoms can be located with accuracy by experiments at the appropriate wavelengths. This positional identification allows for a reference to determine phases for the total reflections, allowing for structure solution from a single crystal (Hendrickson 1997).

MAD data collection strategy was based on the recommendations outlined by a series of papers by Ana Gonzalez (Gonzalez, Pedelacq et al. 1999; Gonzalez 2003; Gonzalez 2003). In short, crystals were first screened to look for reasonable diffraction quality. A fluorescence scan was then taken of the crystal of interest, and f' and f'' were calculated from the scan (see Figure 5-4 for a graphical representation of fluorescence data, f' , and f''). Datasets were then collected at f' , and the decay of the crystal was evaluated before collecting f'' or f_{remote} . In all cases but one, the

decay of the crystals was so significant that collecting more than one wavelength on each crystal was not feasible.

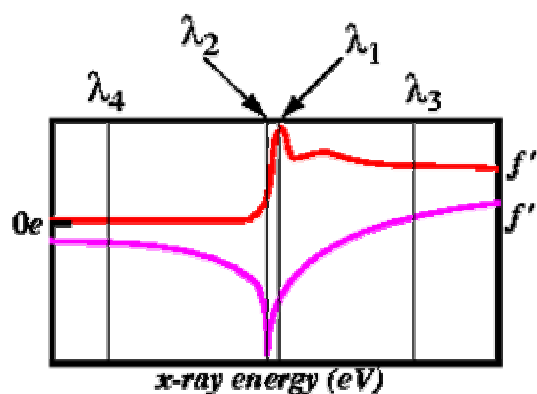


Figure 5-4. Theoretical absorption of selenium in a seleno-methionine experiment. The absorption values are shown along the Y axis, and x-ray energy in eV along the X axis. For a traditional MAD experiment, the researcher would collect data sets at energy λ_1 , then λ_2 , and λ_3 or λ_4 , depending upon the radiation sensitivity of the crystal. However, current literature indicates the necessity of collecting only λ_1 , or perhaps λ_1 and λ_2 , in order to successfully obtain enough anomalous signal for structure determination (Gonzalez 2003). For the data sets used in the current analysis, only data obtained at λ_1 was necessary for full structure solution. Picture taken from (http://www.bmsc.washington.edu/scatter/AS_index.html)

The programs SOLVE and RESOLVE (Terwilliger and Berendzen 1999; Terwilliger 2000) were used to calculate the initial selenomethionine sites and initial phases from SAD data. SOLVE is a highly automated program designed to take the data output from HKL2000, calculate the Patterson and find the heavy atom or selenomethionine peaks, use the input symmetry of the crystal to cross-validate heavy atom sites, and calculate initial phase information utilizing the highest set of heavy atom peaks. If SOLVE is able to find a significant number of peaks, thereby

leading to phases with a high figure of merit, the resultant electron density should be interpretable in many cases (Terwilliger and Berendzen 1999). However, vast improvements in map quality can often be obtained by following SOLVE with RESOLVE. RESOLVE utilizes a maximum likelihood approach to apply density modification to the experimental phases derived from SOLVE (Terwilliger 2000). Information about the solvent boundary, non-crystallographic symmetry, and partial model information (from initial models or molecular replacement solutions) can be utilized in RESOLVE to improve the probability target for the experimental phases as well, thereby leading to a very powerful (and almost completely automated) approach to phase improvement.

Native Data Collection and Data Reduction

Native data collection was carried out at BM-19 and ID-19 at the APS, and at 8.2.1 and 8.2.2 at the ALS. Data indexing, integration, and scaling were carried out using HKL2000 (Minor 1997). After initial models had been built from the SAD data sets, molecular replacement was used to further refine the model in the higher resolution native data sets for each nucleotide. In all cases, molecular replacement was out carried out using AMoRE as outlines in Chapter Three.

Model Building and Refinement

Model building and refinement were carried out in O (Jones, Zou et al. 1991). Small molecule torsion angle and bond length restraints were generated using HIC-UP for MANT nucleotides and for BODIPY molecules (Kleywegt and Jones 1998).

RESULTS AND DISCUSSION

Expression, Purification, and Crystallization of Unlabeled and Selenomethionine

Labeled Untagged $G\alpha_{i1}$

Original rounds of purification of unlabeled $G\alpha_{i1}$ resulted in a pure protein that was slightly larger than expected (43 kDa instead of 40.3 kDa). This result was verified by SDS-PAGE and mass spectra (Figures 5-5 and 5-6). However, this protein crystallized under the conditions outlined for $G\alpha_{i1}$, and also reacted with anti- $G\alpha_{i1}$ antibody. Crystals obtained with untagged protein appeared in the time frame expected for $G\alpha_{i1}$ (overnight, with terminal growth in one week). The crystals also appeared in the published condition to $G\alpha_{i1}$ protein, which uses a relatively uncommon precipitant (ammonium sulfite). Although many proteins crystallize in conditions containing ammonium sulfate, it was not known at the time whether proteins that normally crystallize in ammonium sulfate would also crystallize in ammonium sulfite. It was observed that crystals resulting from certain nucleotide exchanges (especially BODIPY-GTP γ S and BODIPY-GDP) were larger and had better morphology than others (especially MANT-GDP). This led to speculation that

fluorophore nucleotide incorporation into the untagged protein was leading to differential nucleation and growth compared to unlabeled nucleotide, and that the differences between the crystals were due to different affinities of the labeled nucleotides.

The crystals obtained with this protein were not a space group consistent with known $G\alpha_{i1}$ structures. The protein crystallized in space group $P2_12_12$, as opposed to $P3_121$ for activated $G\alpha_{i1}$ crystal forms and I_4 for GDP-bound $G\alpha_i$ (Coleman, Lee et al. 1994). Crystals were harvested, cryoprotected in glycerol, and data sets were collected at ALS crystals. All attempts at solving these structures by molecular replacement failed, indicating that the protein in the crystals were significantly different in conformation from previous $G\alpha_{i1}$ structures. It was possible that the incorporation of fluorophore nucleotides were responsible for these results; in order to find out, the structure could only be solved by utilizing selenomethionine-labeled protein and MAD phasing. Protein produced under selenomethionine growth conditions showed similar purification properties as unlabeled protein, and mass spectra confirmed selenomethionine incorporation. Selenomethionine labeled crystals were grown under identical conditions of nucleotide and precipitant as unlabeled protein (Figure 5-7). The selenomethionine crystals were harvested and cryoprotected in an identical way to the native crystals, and SAD data sets were collected at ALS.

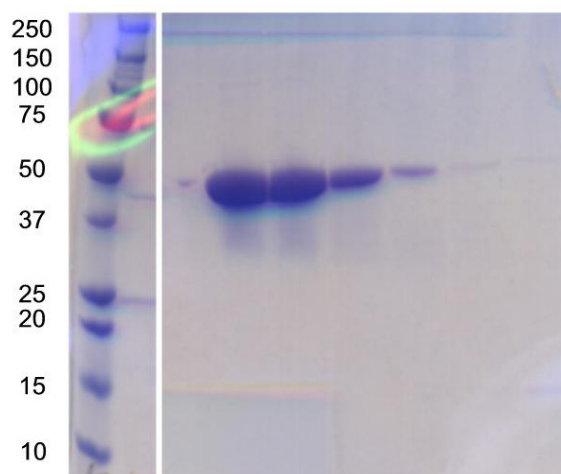


Figure 5-5. SDS-PAGE analysis of gel filtered protein from $G\alpha_{i1}$ preps. This gel shows a typical gel filtration run on the resultant protein from unlabeled untagged $G\alpha_{i1}$ preps. Lanes are 5 μ L of 0.5 mL fractions.

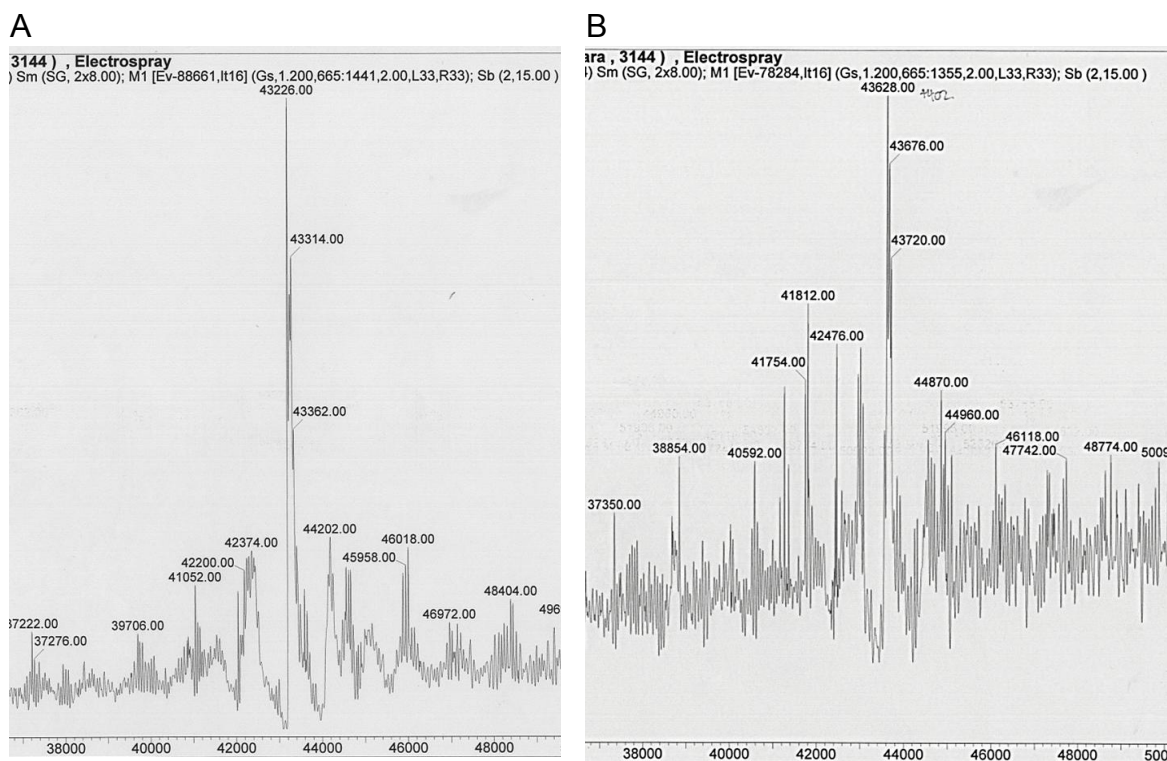


Figure 5-6. Mass spectrometry on unlabeled and selenomethionine labeled protein from $G\alpha_{i1}$ preps. The expected mass of $G\alpha_{i1}$ is 40.3 kDa. The unlabeled protein (A) indicated a larger mass (43.2 kDa), and the selenomethionine labeled sample (B) indicated nine residues were labeled in the protein (402 Dalton difference in mass/46.9 (difference in mass between methionine and selenomethionine) = 8.5).

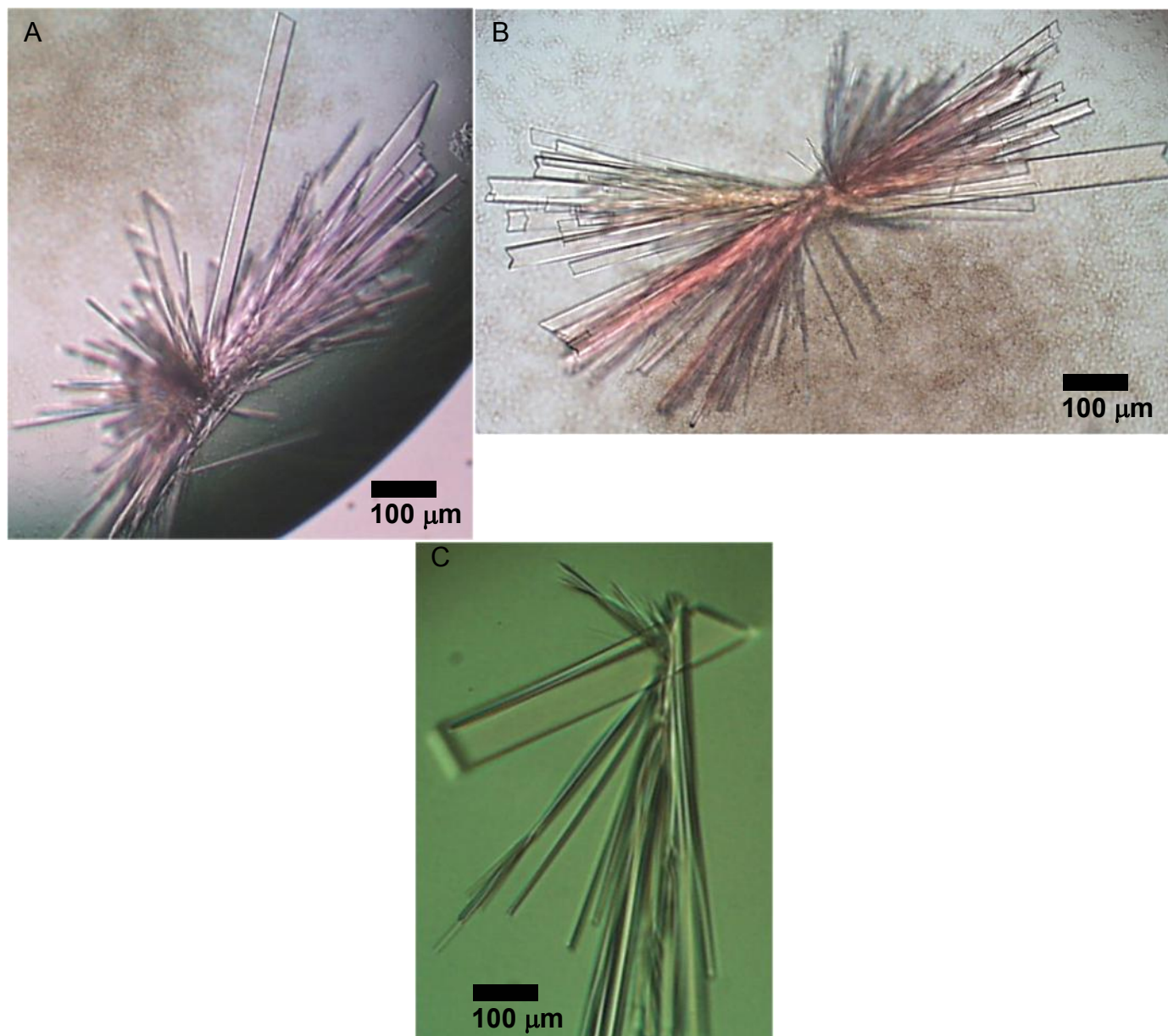


Figure 5-7. Unlabeled and selenomethionine crystals from untagged $G\alpha_{i1}$ preps. Typical crystals from unlabeled preps (A, B) and selenomethionine labeled preps (C) are shown.

The programs SOLVE and RESOLVE were used to find the selenomethionine labeled sites, and a high quality electron density map was calculated using the

phases from the selenium scattering. Upon manual inspection of the density map, secondary structural features of the ras-like domain could be found almost immediately. However, it soon became clear that there were β -barrel structures in the density that showed no similarity to the α -helical domain of $G\alpha_{i1}$, and within the Ras domain density the position of the selenomethionines did not match with the distribution of methionines within the sequence of $G\alpha_{i1}$. Therefore, the density was fitted with a manually drawn polyalanine model, and then the model was submitted to the server DALI (Holm and Sander 1993). This server uses alignment of distance matrices to identify identical folds and domain structures between proteins. The results of DALI were striking (Table 5-2); the model obtained from the selenomethionine map was identical to the protein Ef-Tu from *e. coli*, the source organism from the $G\alpha_i$ preps.

PDB	Z score	RMSD	Description
1EFC	32.3	0.9	elongation factor (eftu) biological_unit
1DAR	16.5	2.3	elongation factor g (ef-g)
1KK1	16.3	3.2	eif2gamma Mutant
1CTQ	13.9	2.5	transforming protein p21H-RAS-1 fragment
1CIP	13.9	2.7	guanine nucleotide-binding protein alpha-1 subunit (gia)
1MKY	13.7	2.5	probable gtp-binding protein enga fragment
1UDX	13.6	4.0	the gtp-binding protein obg
1H65	12.4	2.9	chloroplast outer envelope protein oep34 (toc34, gtp-bi)
1G7R	12.2	3.0	translation initiation factor if2EIF5B fragment
1BYU	12.0	2.5	gtp-binding protein ran fragment (tc4)

Table 5-2. Output of DALI server. The top ten matches in the DALI database are shown. There were 4990 protein chains in the database at the time of this search; a Z-score of more than 2 is considered significant for this output. The RMSD measurement is between the manually built polyalanine model shown in Figure 5-10 and the test protein whose PDB identifier is shown in column 1.

Identification of EF-Tu as the Crystallization Contaminant

The DALI server indicated that the protein crystallized was actually *e. coli* Ef-Tu. To verify this hypothesis, structures of Ef-Tu bound to GDP (1EFC) (Song, Parsons et al. 1999) and GMPPNP (1EXM) (Kjeldgaard, Nissen et al. 1993) were utilized as molecular replacement search models using AMoRE. Initial results indicated clearly that all data sets contained the structure of Ef-Tu in the basal, GDP-bound form. Data sets for each nucleotide tested were compared for resolution and scaling statistics and the best data set for each nucleotide form was subjected to further rounds of refinement.

Data set	(b)GTP γ s	(b)GDP	(sm)(m)GMPPNP	(b)GMPPNP	(m)GDP
Data Collection					
Space group	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
Unit cell					
a	243.691	243.690	244.342	244.112	243.673
b	61.432	61.468	61.416	61.659	61.403
c	67.349	67.325	67.383	67.534	67.263
$\alpha=\beta=\gamma$	90	90	90	90	90
D _{min}	2.3	2.35	2.4	2.7	2.9
Unique					
Reflections	40692	37238	39314	26020	21964
Redundancy	3.2 (1.7)	3.1 (2.2)	17.1 (10.3)	5.5 (4.0)	3.1 (2.1)
Completeness	88.6 (44.4)	86.2 (53.3)	96.6 (86.6)	90.2 (67.8)	93.5 (81.2)
<I/ σ >	9.9 (1.2)	9.4 (1.4)	16.4 (1.9)	8.3 (1.6)	6.9 (1.2)
R _{sym}	10.4 (50.5)	12.1 (44.7)	14.7 (81.2)	16.0 (64.9)	15.2 (60.2)
Mosaicity	0.5	0.3	0.2	0.45	0.4
Wilson B factor	36.61	37.67	47.53	45.69	53.04
Refinement (all statistics are for the GDP:Mg ⁺⁺ bound form with H ₂ O added)					
Resolution	50-2.3	50-2.35	50-2.4	50-2.7	50-2.9
Number of Atoms					
Protein	11936	11936	11936	11936	11936
Water	201	174	121	148	90
Other (nucleotide,ions)	88	88	88	88	88
RMS deviations					
Bond lengths	1.22	1.22	1.28	1.29	1.28
Bond angles	0.007	0.005	0.007	0.008	0.006
RMS B factors					
Bonded main chain	1.45	2.4	2.6	1.25	2.04
Bonded side chain	2.15	3.8	3.8	1.63	3.05
Average B factor	35.62	35.38	42.01	31.59	36.23
R _{work}	22.87	22.66	21.35	21.71	21.62
R _{free}	26.11	26.57	25.99	26.45	27.65

Table 5-3. Description of the data sets used for structure determination. Each column indicates a separate data set used to solve a structure. The first row indicates which nucleotide should have been in the structure and is abbreviated as follows: (b) = BODIPY, (m) = MANT. (sm) = selenomethionine; all other data sets were native and were solved using molecular replacement methods. The data sets are arranged from left to right in order of decreasing resolution.

Structural Analysis of EF-Tu

The results from molecular replacement and from selenomethionine labeled samples indicated that the protein crystallized was Ef-Tu from *e. coli*. However, since Ef-Tu is also a G protein, structural analysis was undertaken to discern whether the proteins crystallized in an active or inactive conformation. It would be expected that under the conditions used for nucleotide exchange, a certain percentage of protein might have crystallized with the proper fluorophore nucleotide, perhaps undergoing a conformational change in the process.

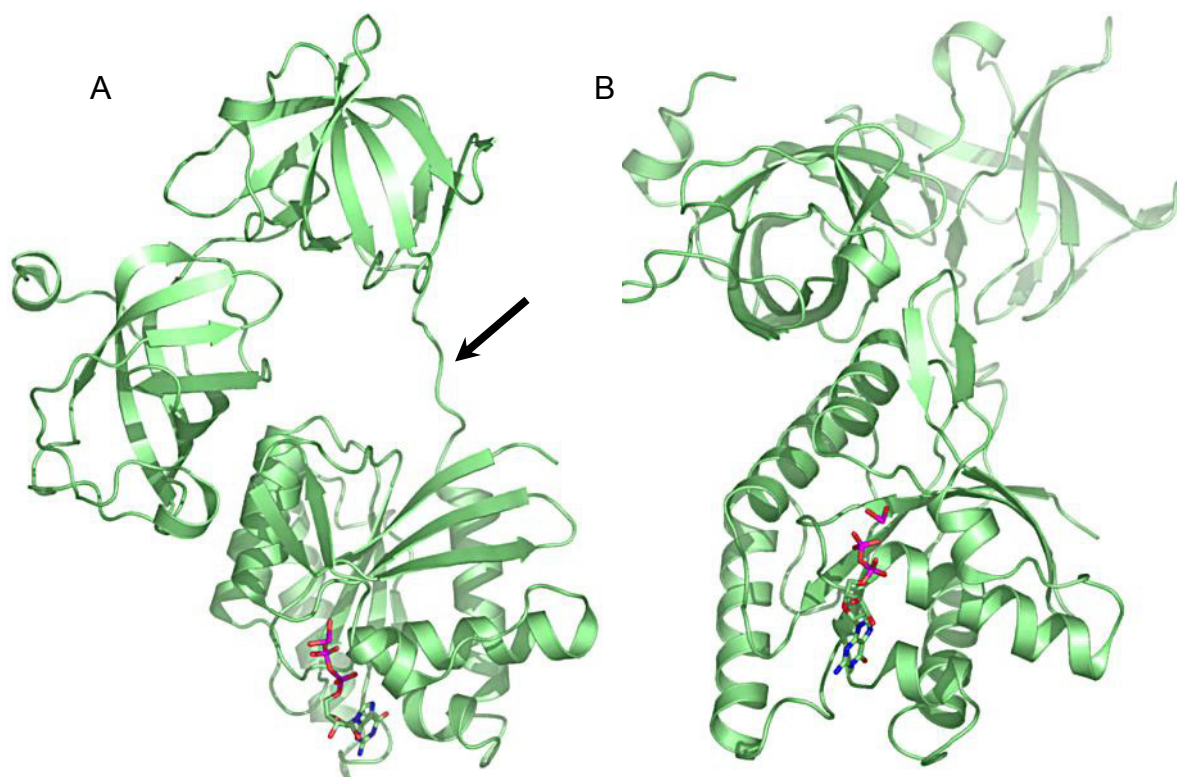


Figure 5-8. Structure of Ef-Tu in the GDP-bound conformation. The two views are related by approximately $+90^\circ$. The protein is composed of three domains: the ras-like domain, shown bound to GDP, and two β barrels. The largest conformational change that occurs when GTP binds to Ef-Tu is a domain movement

about the linker shown with an arrow in (A), which brings the β barrels in close proximity to the ras-like domain. This change is shown in Figure 5-9.

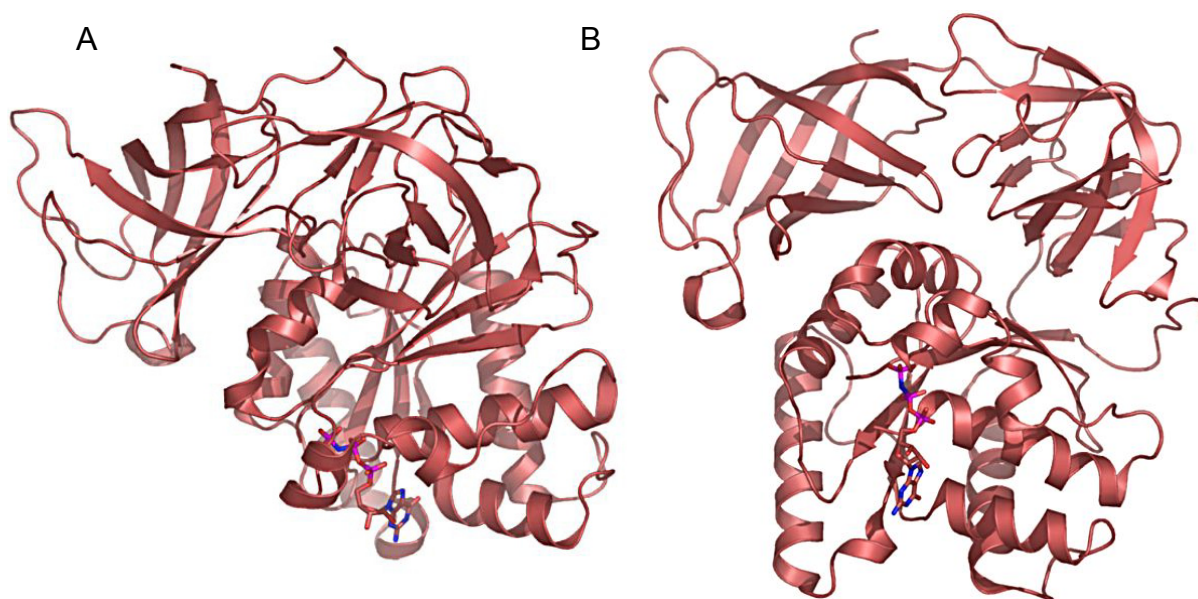


Figure 5-9. Structure of Ef-Tu in an active conformation. Ef-Tu bound to GMPPNP is shown in approximately the same orientation as Figure 5-8. In (A), the β -barrel domains can be described as having collapsed onto the ras-like domain, making the overall fold of the protein quite distinguishable from the same orientation in Figure 5-8 (A). In (B), the Ras domain can be seen to be quite similar to the form in the GDP-bound conformation, but the arrangement of the β barrel domain is rotated and packed against the Ras domain.

The difference between Ef-Tu in the GDP (Figure 5-8) and active conformations (Figure 5-9) is quite dramatic compared to other G proteins. The switch I region undergoes a structural rearrangement from β strand to a helix in the active form (Song, Parsons et al. 1999). The switch II helix loses a helical turn at its N terminus and gains a helical turn at its C terminus, and changes its orientation by about 40° (Song, Parsons et al. 1999). Large changes can be seen between the Ras domain and the β barrel domains, as shown in Figures 5-8 and 5-9. The interactions

between the two β barrel domains are preserved, but the Ras domain rotates about 90° relative to the β barrel domains, closing the distance between them by about 10Å.



Figure 5-10. The structure of the G α_{i1} contaminant. After ascertaining the nature of the β barrels connected to the Ras domain, a manual polyaniline model was built into the electron density obtained through selenomethionine phasing. After multiple rounds of building and refinement, this polyaniline model for one of the molecules in the asymmetric unit was obtained. This model was used as an input to the DALI server to aid in identification.

Analysis of the model built from selenomethionine data (Figure 5-10), along with models from the molecular replacement solutions (Figure 5-16), confirm that all of the proteins crystallized in the current study are in the basal conformation (GDP-bound).

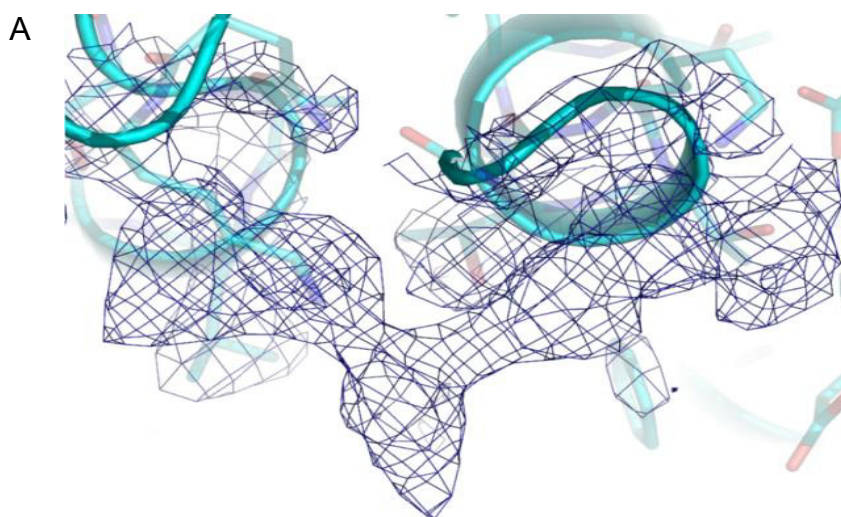
State of the Bound Nucleotide in each EF-Tu Structure

After determining the global conformations of Ef-Tu in the contaminant crystals, it was reasonable to ask what the nucleotide state of each type of data set was. Since Ef-Tu is a G protein, which undergoes cycles of exchange and hydrolysis, it was important to make sure that none of the structures obtained actually contained the desired fluorophore nucleotide. Even at low occupancy, it could be possible that the proteins had crystallized with the proper fluorophore nucleotide, even if the protein had not adopted the proper activated conformation. To accomplish this analysis, molecular replacement with the basal conformation of Ef-Tu as a search model was used to find initial density for each of the five nucleotide variables. The nucleotide and any water or magnesium ions were removed from the search model, such that the initial electron density obtained was unbiased in terms of nucleotide phasing. In each case, the nucleotide binding pocket was manually analyzed for the presence of absence of nucleotide (Figure 5-11). In each case studied, positive density was observed in the nucleotide binding pocket, indicating the presence of nucleotide of some variety. Next, a model of GDP was added to the protein model, and one cycle of refinement carried out in CNS. The density was again examined, and for each

data set additional positive density was still present. However, it became obvious at this stage that the positive density left could not be explained by modeling in the structure of either BODIPY or MANT, but instead was due to the presence of magnesium ion and ordered waters in the nucleotide binding site (see Figures 5-11 through 5-15). Another round of refinement with GDP, magnesium ion, and water was carried out, and the resulting density analyzed. See Figures 5-11 through 5-15 below for a graphical account of this process and the final round of refinement density for each of the five data sets studied.

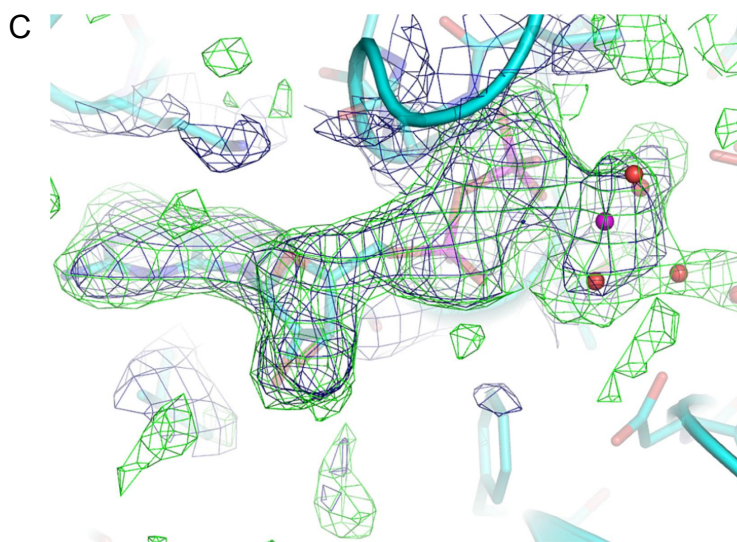
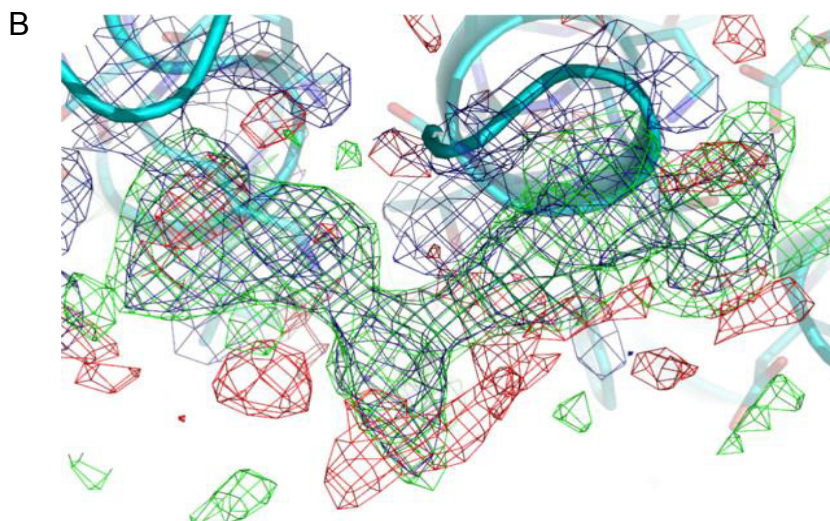
Figure 5-11. The process of building nucleotide models into electron density.

In panels (A) – (D), the process of building a nucleotide model for the BODIPY-GDP data set is outlined. These structures are from a crystal in which the protein has gone through exchange for BODIPY-GDP. The views in (A), (B), and (D) are looking onto the protein in the nucleotide binding pocket, and the view in (C) is roughly +90°. In (A), the 2Fo-Fc density for the nucleotide is clearly seen in blue (contoured at 1.5 σ), indicating the presence of nucleotide in the binding pocket.



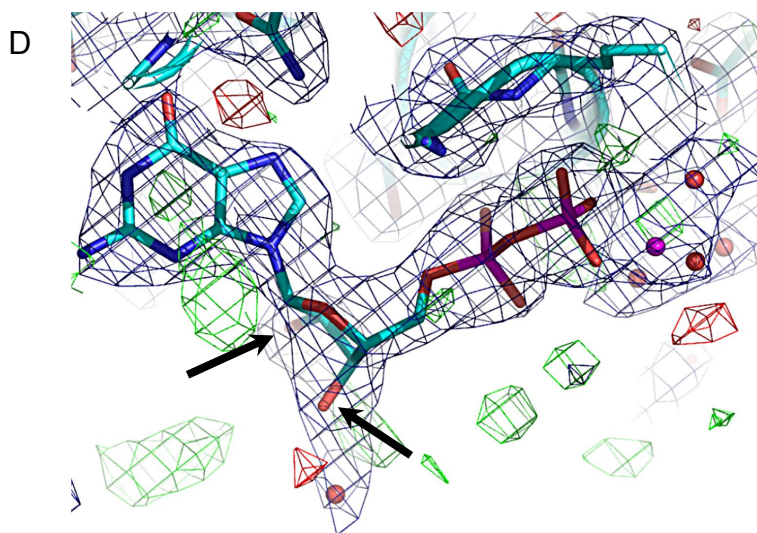
In (B), Fo-Fc density is displayed at a contour level of +2.5 σ (green) and -2.5 σ (red) in order to build nucleotide into the observed density for this model. In addition to

the observed 2Fo-Fc density, the “positive” density indicates areas where molecules should be added to the model to contribute to the density. In contrast, the “negative” density indicates areas where modeling should not be done. In practice, visualizing these three types of density maps allows the crystallographer to build a more accurate model for further refinement.



(C) Based on previously published coordinates for Ef-Tu, a molecule of GDP, a magnesium ion, and several water molecules are fit into the observed 2Fo-Fc and positive Fo-Fc density for the BODIPY-GDP data set. This view is slightly tilted relative to the previous figures, such that the guanine ring is almost horizontally planar. Negative density is not shown for the sake of clarity.

Once a reasonable model is built into the density, the new model is subjected to a round of refinement and a new map is built from the phases of the new model. The resultant map is then compared with the old map to see if changes in the model can be correlated with improvements in the map quality and the disappearance of positive density.



(D) The final round of refinement for the model representing the BODIPY-GDP data set. Notice the preponderance of 2Fo-Fc density surrounding the model, and the relative disappearance of positive and negative density in the nucleotide binding pocket. In this picture, the 2' and 3' oxygens of the ribose moiety are indicated with arrows – the lack of coherent positive density in these regions indicates that the BODIPY fluorophore is not attached to the GDP molecule in this structure.

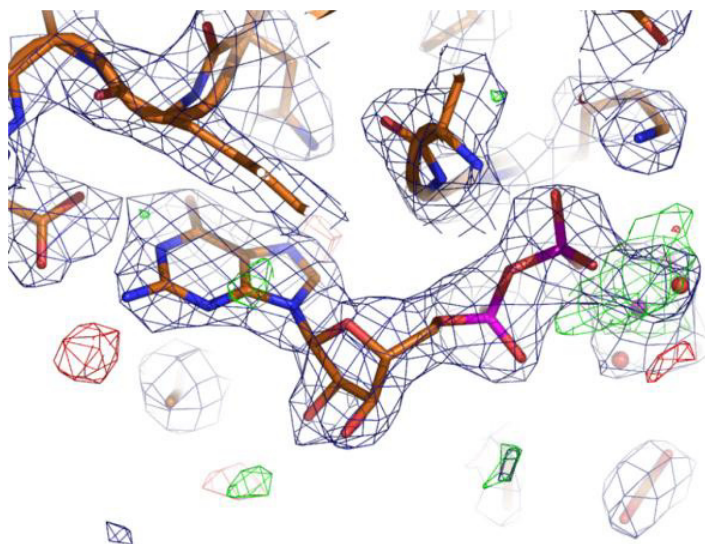


Figure 5-12. The final model for the MANT-GDP data set. This model contains a GDP molecule, one magnesium ion, and water molecules in the nucleotide binding pocket. There is no clear density connected to the 2' or 3' oxygen of the ribose group, and therefore no evidence to support the addition of a MANT molecule to the nucleotide model.

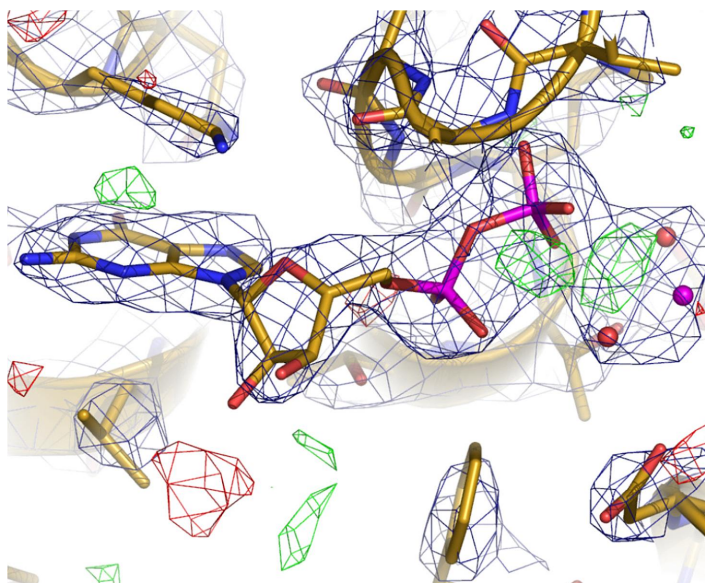


Figure 5-13. The final model with refined density for the BODIPY-GMPPNP data set. There is slight positive density around the magnesium ion, but there is no evidence to support building a γ phosphate group in this region. There is also no evidence to support the addition of a BODIPY moiety in the region of either the 2' or 3' oxygen of the ribose ring.

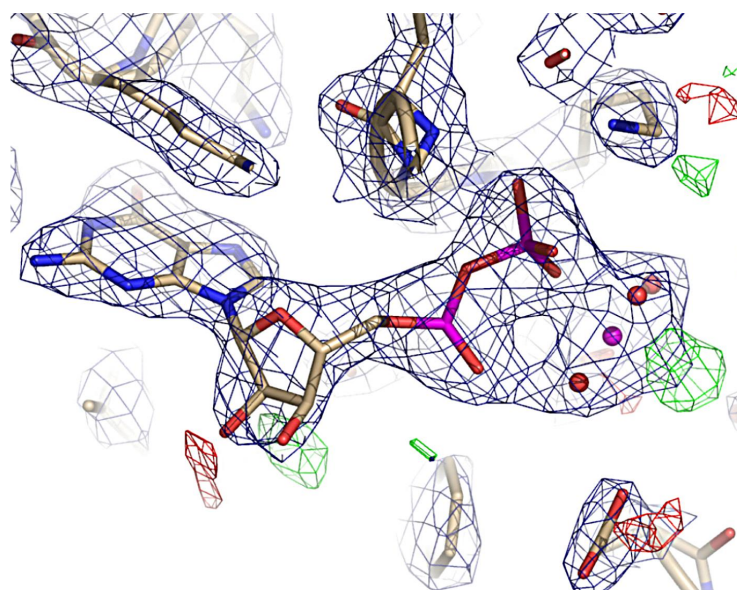


Figure 5-14. Final density surrounding the MANT-GMPPNP data set. There is no evidence supporting the inclusion of a γ phosphate or MANT fluorophore chemical group into this density.

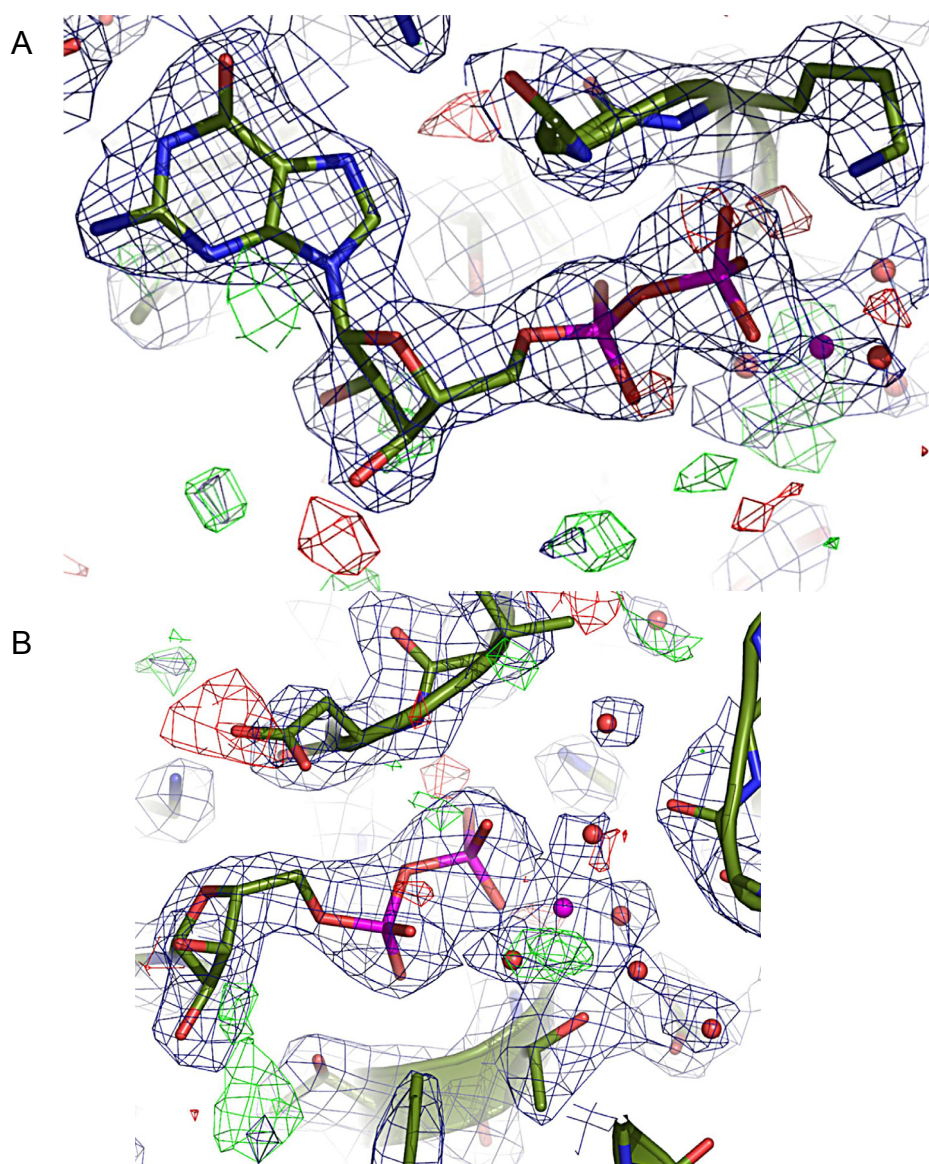


Figure 5-15. Final model of the data set containing the BODIPY-GTP γ S nucleotide. In this case, the BODIPY chemistry should be connected through the sulfur contained within the γ phosphate. The first view shows the quality of the density surrounding the guanine ring and sugar of the nucleotide. The second view is centered upon the α and β phosphates of the modeled GDP. There is no evidence in this binding pocket for either a γ phosphate or a BODIPY moiety connected to that phosphate.

Analysis of the Conformational State of each EF-Tu Structure

Refinement of the five nucleotide types was completed and the structures were superimposed upon each other. In conclusion, all five structures are clearly in the same conformational state (that of the GDP-bound or basal conformation of Ef-Tu), and analysis of the nucleotide binding sites indicate no exchange for fluorophore nucleotide has taken place in any of the five types of structures studied.

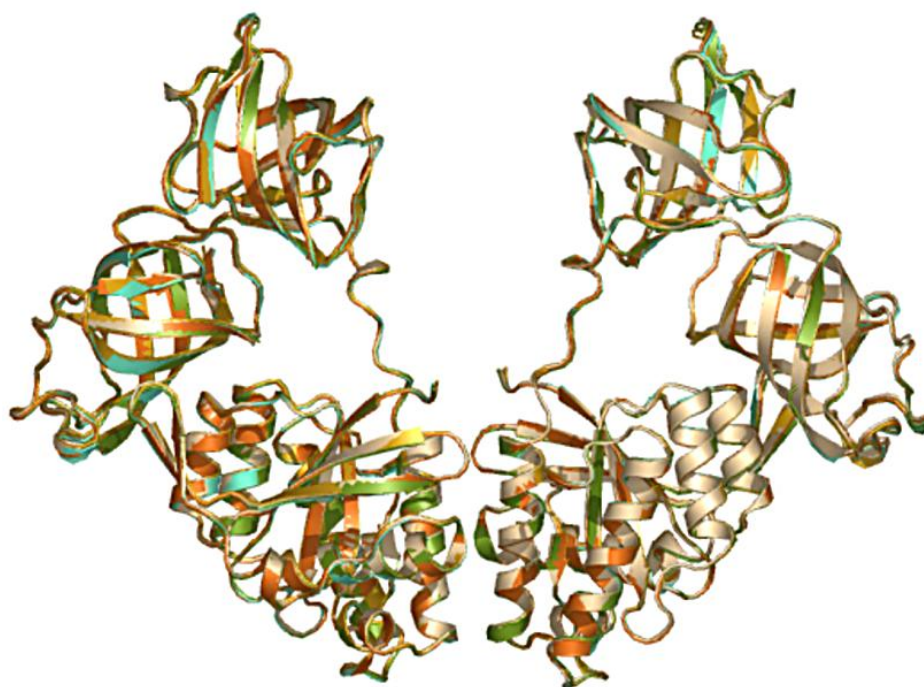


Figure 5-16. Superposition of the five nucleotide data sets collected. The color scheme is identical to Figures 5-11 through 5-15. Compare these two molecules in the asymmetric unit to the structures of the GDP and GMPPNP forms of Ef-Tu shown in Figures 5-8 and 5-9.

Troubleshooting Strategies for the Second Attempt at Crystallization

It was determined that the cause of the incorrect purification procedure was due mainly to a problem in the expression conditions used for the untagged $G\alpha_i$ prep.

For all preps after the first, a key ingredient in the induction protocol had been omitted; chloramphenicol at 1 $\mu\text{g/L}$ of cells is normally added to boost the production of $G\alpha_i$ protein, and this step not been transferred into the protocol used for protein production. Trials of cells where chloramphenicol was added led to overexpression of $G\alpha_{i1}$, which then crystallized under standard conditions with different crystal morphology than that of Ef-Tu. However, the entire project could be streamlined if affinity chromatography could be utilized to produce $G\alpha_{i1}$, and so attempts to express and purify $G\alpha_{i1}$ as a GST fusion were undertaken. It was also decided that during the exchange and crystallization protocols a considered effort would be undertaken to keep the levels of fluorophore nucleotide around 10x molar excess over protein concentration, in order to encourage exchange and equilibrium binding of fluorophore nucleotide. The results of these changes are outlined in Chapter Six.

BIBLIOGRAPHY

- Bokoch, G. M., T. Katada, et al. (1984). "Purification and properties of the inhibitory guanine nucleotide-binding regulatory component of adenylate cyclase." J Biol Chem **259**(6): 3560-7.
- Coleman, D. E., A. M. Berghuis, et al. (1994). "Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis." Science **265**(5177): 1405-12.
- Coleman, D. E., E. Lee, et al. (1994). "Crystallization and preliminary crystallographic studies of Gi alpha 1 and mutants of Gi alpha 1 in the GTP and GDP-bound states." J Mol Biol **238**(4): 630-4.
- Coleman, D. E. and S. R. Sprang (1999). "Structure of Gialpha1.GppNHp, autoinhibition in a galpha protein-substrate complex." J Biol Chem **274**(24): 16669-72.
- Doublie, S. (1997). Preparation of Selenomethionyl Proteins for Phase Determination. New York, Academic Press.
- Gille, A. and R. Seifert (2003). "2'(3')-O-(N-methylanthraniloyl)-substituted GTP analogs: a novel class of potent competitive adenylyl cyclase inhibitors." J Biol Chem **278**(15): 12672-9.
- Gille, A. and R. Seifert (2003). "Low-affinity interactions of BODIPY-FL-GTPgammaS and BODIPY-FL-GppNHp with G(i)- and G(s)-proteins." Naunyn Schmiedebergs Arch Pharmacol **368**(3): 210-5.
- Gille, A. and R. Seifert (2003). "MANT-substituted guanine nucleotides: a novel class of potent adenylyl cyclase inhibitors." Life Sci **74**(2-3): 271-9.
- Gonzalez, A. (2003). "Faster data-collection strategies for structure determination using anomalous dispersion." Acta Crystallogr D Biol Crystallogr **59**(Pt 2): 315-22.
- Gonzalez, A. (2003). "Optimizing data collection for structure determination." Acta Crystallogr D Biol Crystallogr **59**(Pt 11): 1935-42.
- Gonzalez, A., J. Pedelacq, et al. (1999). "Two-wavelength MAD phasing: in search of the optimal choice of wavelengths." Acta Crystallogr D Biol Crystallogr **55** (Pt 8): 1449-58.

- Hendrickson, W. A. (1991). "Determination of macromolecular structures from anomalous diffraction of synchrotron radiation." Science **254**(5028): 51-8.
- Hendrickson, W. A. a. O., C. M. (1997). Phase Determination from Multiwavelength Anomalous Diffraction Measurements. New York, Academic Press.
- Holm, L. and C. Sander (1993). "Protein structure comparison by alignment of distance matrices." J Mol Biol **233**(1): 123-38.
- Jameson, E. E., J. M. Cunliffe, et al. (2003). "Detection of G proteins by affinity probe capillary electrophoresis using a fluorescently labeled GTP analogue." Anal Chem **75**(16): 4297-304.
- Jones, T. A., J.-Y. Zou, et al. (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models." Acta Crystallographica Section A **47**(2): 110-119.
- Kjeldgaard, M., P. Nissen, et al. (1993). "The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation." Structure **1**(1): 35-50.
- Kleywegt, G. J. and T. A. Jones (1998). "Databases in protein crystallography." Acta Crystallogr D Biol Crystallogr **54**(Pt 6 Pt 1): 1119-31.
- Lee, E., M. E. Linder, et al. (1994). "Expression of G-protein alpha subunits in *Escherichia coli*." Methods Enzymol **237**: 146-64.
- McEwen, D. P., K. R. Gee, et al. (2001). "Fluorescent BODIPY-GTP analogs: real-time measurement of nucleotide binding to G proteins." Anal Biochem **291**(1): 109-17.
- Minor, Z. O. a. W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. Methods in Enzymology. (New York), Academic Press. **276**: 307-326.
- Mixon, M. B., E. Lee, et al. (1995). "Tertiary and quaternary structural changes in Gi alpha 1 induced by GTP hydrolysis." Science **270**(5238): 954-60.
- Remmers, A. E. (1998). "Detection and quantitation of heterotrimeric G proteins by fluorescence resonance energy transfer." Anal Biochem **257**(1): 89-94.
- Remmers, A. E., C. Engel, et al. (1999). "Interdomain interactions regulate GDP release from heterotrimeric G proteins." Biochemistry **38**(42): 13795-800.

- Remmers, A. E. and R. R. Neubig (1996). "Partial G protein activation by fluorescent guanine nucleotide analogs. Evidence for a triphosphate-bound but inactive state." J Biol Chem **271**(9): 4791-7.
- Remmers, A. E., R. Posner, et al. (1994). "Fluorescent guanine nucleotide analogs and G protein activation." J Biol Chem **269**(19): 13771-8.
- Song, H., M. R. Parsons, et al. (1999). "Crystal structure of intact elongation factor EF-Tu from Escherichia coli in GDP conformation at 2.05 Å resolution." J Mol Biol **285**(3): 1245-56.
- Terwilliger, T. C. (2000). "Maximum-likelihood density modification." Acta Crystallogr D Biol Crystallogr **56** (Pt 8): 965-72.
- Terwilliger, T. C. and J. Berendzen (1999). "Automated MAD and MIR structure solution." Acta Crystallogr D Biol Crystallogr **55** (Pt 4): 849-61.

CHAPTER SIX

Crystallization of $G\alpha_{i1}$ Bound to Fluorophore Nucleotides: Second Attempt

This chapter outlines the progress made in crystallizing $G\alpha_{i1}$ obtained from a GST construct and the second attempt to crystallize $G\alpha_{i1}$ bound to fluorophore labeled nucleotides. A GST construct was made which expressed, could be purified, and crystallized well under standard $G\alpha_{i1}$ conditions. Unfortunately, the crystallization precipitant reacted unfavorably with the fluorophores, and crystals of the $G\alpha_{i1}$ •fluorophore complex were not obtainable.

MATERIALS AND METHODS

Expression and Purification of GST- $G\alpha_{i1}$ Constructs

The GST- $G\alpha_{i1}$ construct was obtained from C. Thomas in the Sprang lab. The construct contained a TEV site between the GST and $G\alpha_i$ (Figure 6-1). After TEV cleavage, three residues remain N-terminal to the initiating methionine of $G\alpha_i$. The GST-mod $G\alpha_{i1}$ construct was obtained from A. Adhikari in the Sprang lab. This construct also contains a TEV site, but after TEV cleavage the first remaining residue is the second residue of $G\alpha_i$ (Figure 6-1). For both constructs, expression was in BL21(DE3) cells grown in LB media at 30°C; at $OD_{600}=0.5$, IPTG was added to a final concentration of 300 μ M and the cells were induced overnight at 30°C. Purification of both constructs used identical protocols: the cells were harvested after overnight induction at 2600g for 30 minutes. Cells were lysed using sonication (5 s

on, 5 s off, 5 minutes total sonication time) and then incubated with DNase I (5 mg +2 mM MgCl₂) and protease inhibitor (1 tablet EDTA-free Complete cocktail, Roche) for 15 minutes at 4°C. The supernatant was clarified by ultracentrifugation for 30 minutes at 125,000g. The clarified supernatant was added to 2-3 mL glutathione resin pre-equilibrated in buffer GST-A (20 mM HEPES pH 8, 5 mM β-ME, 1 mM EDTA), washed with GST-B (GST-A + 0.5M NaCl), and eluted in GST-C (GST-A + 20 mM glutathione). TEV protease was added to the GST eluate and incubated overnight. The resultant protein was loaded onto either a phenyl-superose column pre-equilibrated in buffer PHE-A (20 mM HEPES pH 8, 3 mM EDTA, 5 mM β-ME, and 1.2M (NH₄)SO₄) and eluted with buffer PHE-B (20 mM HEPES pH 8, 3 mM EDTA, 5 mM β-ME), or loaded onto a Hitrap Q column pre-equilibrated in buffer Q-A (20 mM HEPES pH8, 1 mM EDTA, 5 mM β-ME) and eluted in buffer Q-B (Q-A + 0.5 M NaCl). In either case, fractions were analyzed using SDS-PAGE and pooled, then incubated with fresh glutathione resin for at least 1 hour with rocking at 4°C. The resin was then removed from the protein by low-speed centrifugation and the resulting supernatant checked for removal of residual GST by SDS-PAGE. The protein was either concentrated directly for crystallography at this stage, or was passed over a Sephadex 200:75 tandem column for further purification as needed. Yields of GST-Gα_i and GST-modGα_i were usually 15 mg and 10 mg per liter of cell culture, respectively. For both constructs, purified proteins were identified using mass spectrometry and Edman sequencing to verify the proteins as Gα_i.

Modification of Nucleotide Exchange Protocol

Exchange was accomplished by adding 2 mM MgCl_2 and 10x molar excess of the nucleotide to concentrated $\text{G}\alpha_i$. Typical concentrations of protein used in the experiment ranged from 25-250 μM protein. The resulting solution was placed at 30 °C for three hours, and then placed overnight at 4 °C. The protein mixture was then diluted with at least 10 volumes of buffer GF and concentrated to at least 10 mg/mL. The resultant protein solution was spiked with 1 mM EDTA, 2 mM MgCl_2 , and 10x molar excess nucleotide before crystallization trials.

RESULTS AND DISCUSSION

Purification of GST- $\text{G}\alpha_{i1}$ Constructs

The GST- $\text{G}\alpha_{i1}$ construct expressed well and was purified in a straightforward fashion (see Figure 6-1). However, this protein did not crystallize after TEV cleavage, although it was quite pure and the conditions for $\text{G}\alpha_{i1}$ crystallization are well established (Figure 6-2) (Coleman, Berghuis et al. 1994; Coleman, Lee et al. 1994; Mixon, Lee et al. 1995; Coleman and Sprang 1998; Coleman and Sprang 1999). After discussion with members of the lab, it was decided that a new construct should be attempted, in which the residual amino acids left after TEV cleavage were eliminated down to the second residue in the $\text{G}\alpha_{i1}$ protein. This was based on the observation that for untagged $\text{G}\alpha_{i1}$ protein, the initial methionine is not present in

mass spectra or sequencing data. In practice this led to the modG α_{i1} construct, which after TEV cleavage begins with Gly2 of G α_{i1} , as opposed to the Gly(-2)-Ala(-1)-Met1-Gly2 sequence left after TEV cleavage of GST-G α_{i1} .

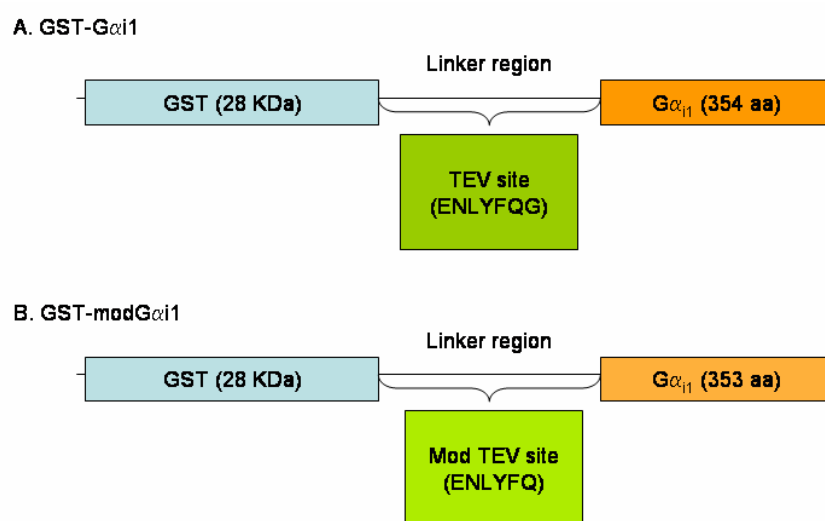


Figure 6-1. Schematic of GST-G α_{i1} and GST-modG α_{i1} constructs.

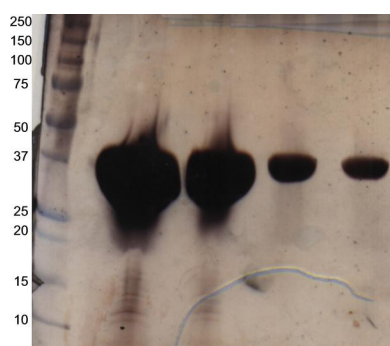


Figure 6-2. Purity of GST-G α_{i1} . After purification, cleavage, and removal of GST, SDS-PAGE was run to verify the purity of the G α_{i1} protein. The gel is silver stained using the Bio-Rad Silver Stain Plus kit; the concentrations of the proteins in each lane are 10 mg/mL, 5 mg/mL, 1 mg/mL, and 0.5 mg/mL. The gel is overloaded to show the absence of residual GST in the preparation.

The expression and purification of modG α_{i1} were quite similar to that of GST-G α_{i1} (see Figures 6-3 through 6-6), and the final purity of modG α_{i1} protein was

comparable to that of GST-G α_{i1} (Figure 6-2 and Figure 6-6). Mass spectra and protein sequencing were used to verify the identity of the modG α_{i1} protein.

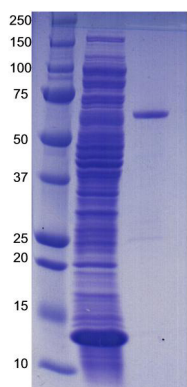


Figure 6-3. GST-modG α_{i1} . The two lanes show the total cell supernatant after clarification using ultracentrifugation and the elution from the GST affinity column. The expected molecular weight of the GST-modG α_{i1} fusion is ~68 kDa.

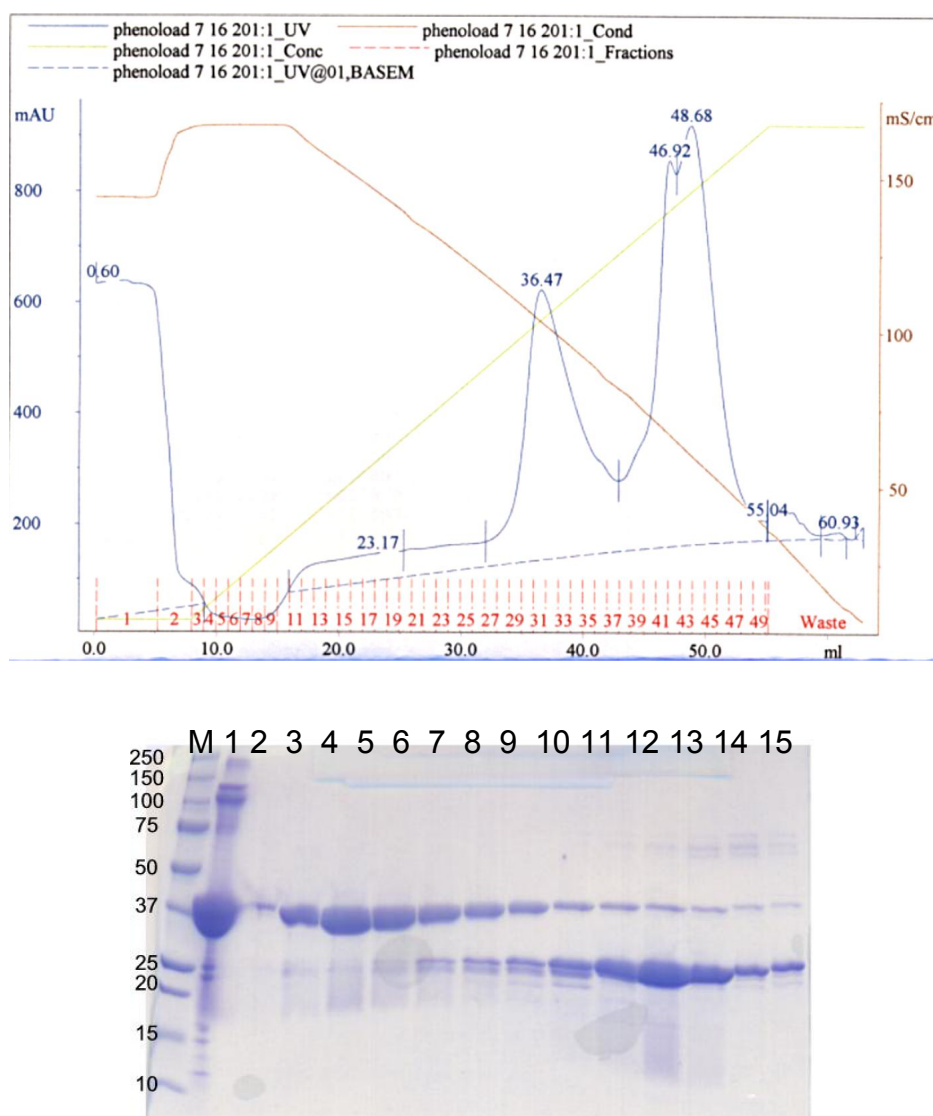


Figure 6-4. Purification by phenyl superose. Fractions from a preparation of GST-mod $G\alpha_{i1}$ run over a phenyl superose column are shown. The first lane shows untagged partially purified $G\alpha_{i1}$ as a (+) control. In this buffer system $G\alpha_{i1}$ elutes first (lanes 2-9), followed by GST (lanes 7-15). Although the peaks on the chromatogram look separated, the purification is only partial and the pooled $G\alpha_{i1}$ from this column must be re-run over glutathione resin to remove residual GST.

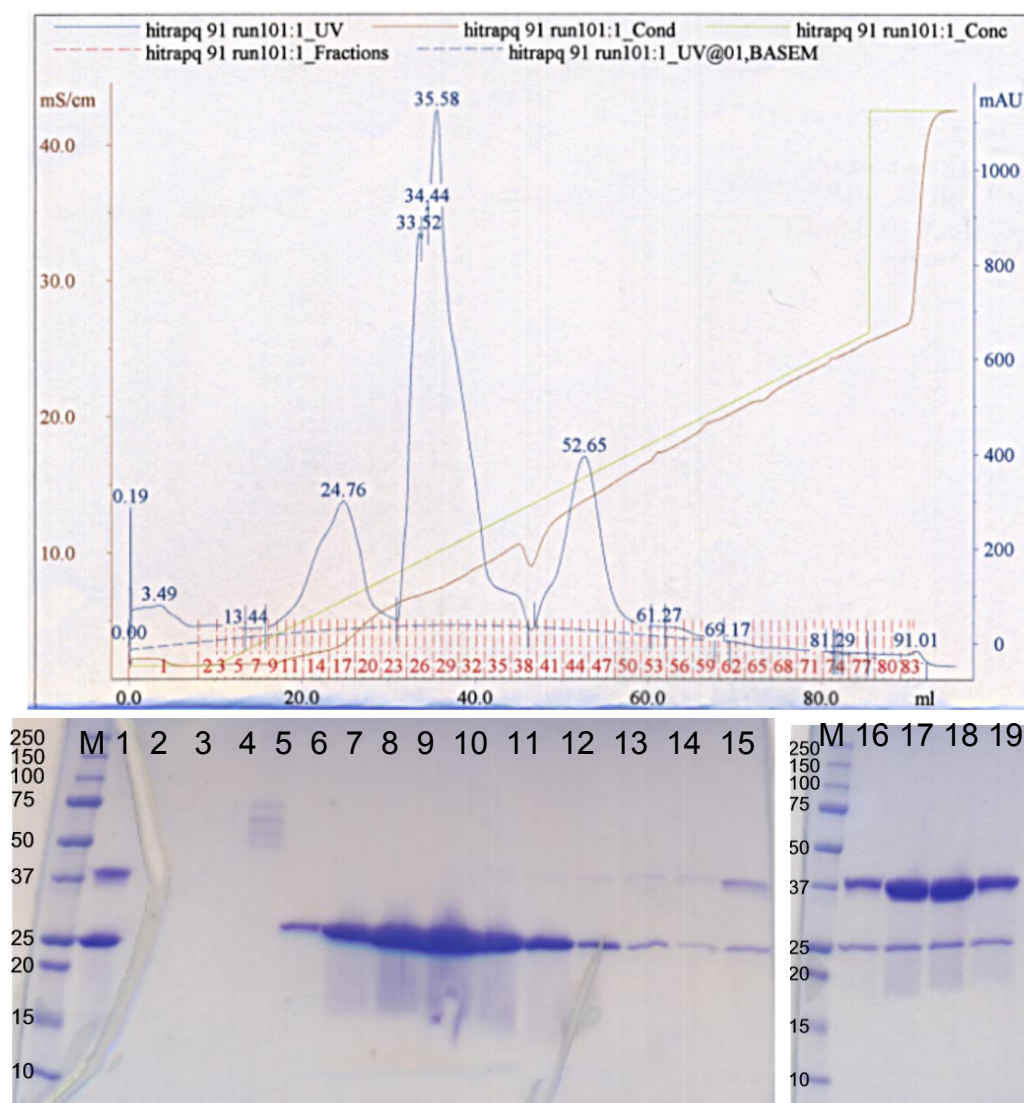


Figure 6-5. Purification using HiTrapQ. After TEV cleavage, the GST-modG α_{i1} protein can be purified utilizing ion exchange chromatography. The first lane contains the protein mixture before loading; spread across two gels are the fractions collected during a gradient elution. Under these buffer conditions, GST elutes first off of the column (lanes 6-13), followed by G α_{i1} (lanes 15-19). The third peak from the chromatogram contains neither GST nor G α_{i1} and is not shown on these gels. Note the residual GST in the G α_{i1} containing fractions; as with phenyl superose, the pooled fractions from the HiTrapQ column must be run again over glutathione resin to ensure purity of the G α_{i1} protein.

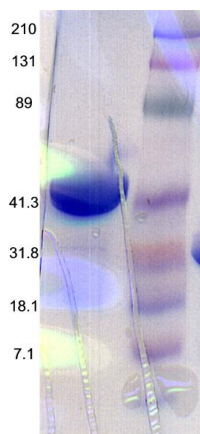


Figure 6-6. Final purity of the GST-modG α_{i1} protein. The purified G α_{i1} is at 10 mg/mL concentration.

Crystallization of GST-G α_{i1} and GST-modG α_{i1}

GST-G α_{i1} was successfully expressed, purified, and set up in hanging drop plates. However, no crystals were obtained under the standard ammonium sulfite conditions for untagged G α_{i1} , even though the purity of the protein was sufficient for crystallography. Surprisingly, the GST-modG α_{i1} construct crystallized quickly under identical conditions to the untagged G α_{i1} protein, and further testing by x-ray diffraction confirms the correct space group and unit cell parameters for both the GDP and GTP γ S forms of the crystals (Figure 6-7).

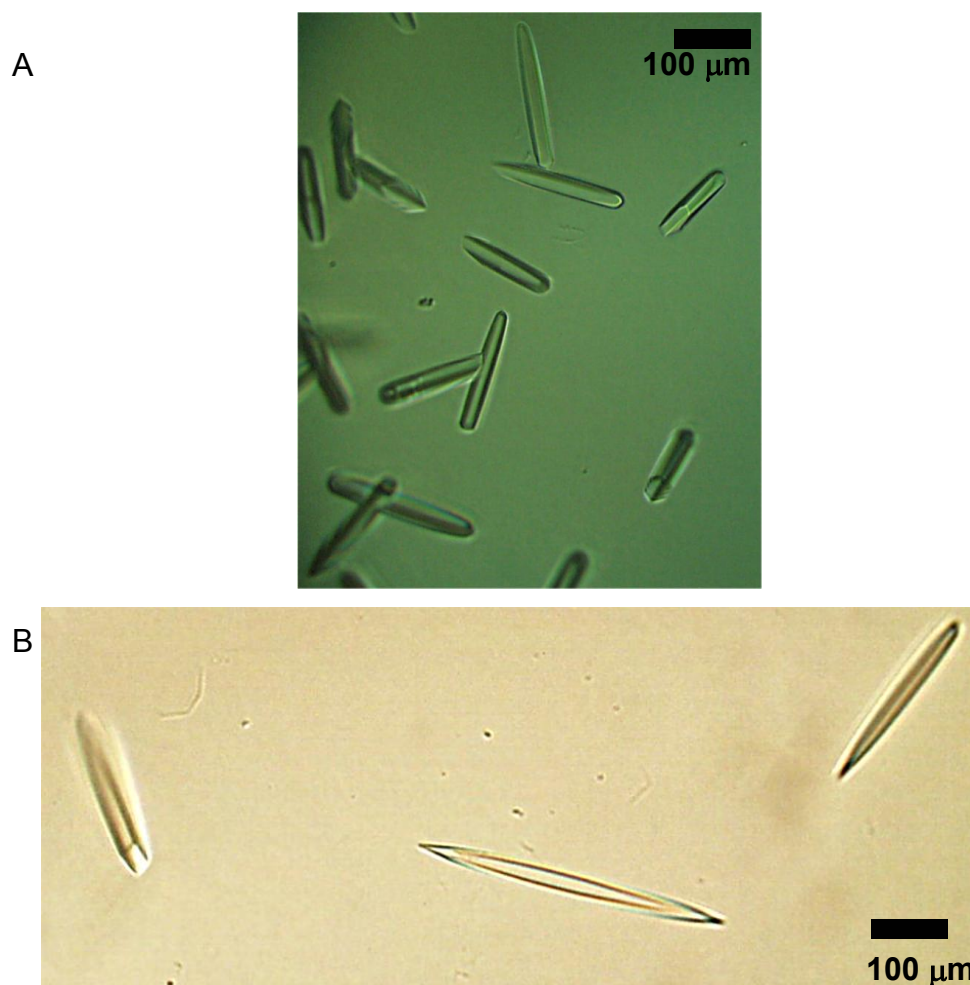


Figure 6-7. GTP γ S and GDP-bound crystals of modG α_{i1} . The activated form of the modG α_i protein (GTP γ S bound) is shown in (A), and GDP-bound modG α_{i1} is shown in (B). The crystals have different morphologies, different space groups, and different cell constants from each other, and so can be readily categorized.

Crystallization of a G α_{i1} •Fluorophore Complex

Exchange of endogenous GDP for fluorescent nucleotide was straightforward using standard techniques. After exchange, either dilution and concentration or gel filtration was used to remove excess nucleotide from solution. There was no

difference seen in the crystallographic experiments between the two methods. However, after nucleotide exchange the protein no longer crystallized under standard ammonium sulfite $G\alpha_{i1}$ conditions. The protein instead precipitated within 24 hours. In addition, the color of the fluorescent nucleotide changed in the case of BODIPY derivatives, from a deep pink color to colorless within 24 hours. This led to the hypothesis that the crystallization condition was affecting the fluorophore nucleotide in a manner which was counterproductive to crystallographic analysis. To test this hypothesis, a series of spectroscopic measurements were undertaken to analyze the effect of crystallization precipitants on the fluorophore-conjugated nucleotides.

Effect of Ammonium Sulfite Crystallization Condition on Fluorophore Nucleotides

In order to test whether ammonium sulfite had a deleterious effect on BODIPY and MANT fluorophores, a series of spectroscopic measurements were carried out. In the simplest set of experiments (Figures 6-8 through 6-11), fluorophore nucleotide solutions were mixed directly with crystallographic conditions (2 M $(\text{NH}_4)\text{SO}_4$, 100 mM Na-Acetate pH 6), and scans taken every 10 minutes to assess the effect of the solution upon the nucleotide.

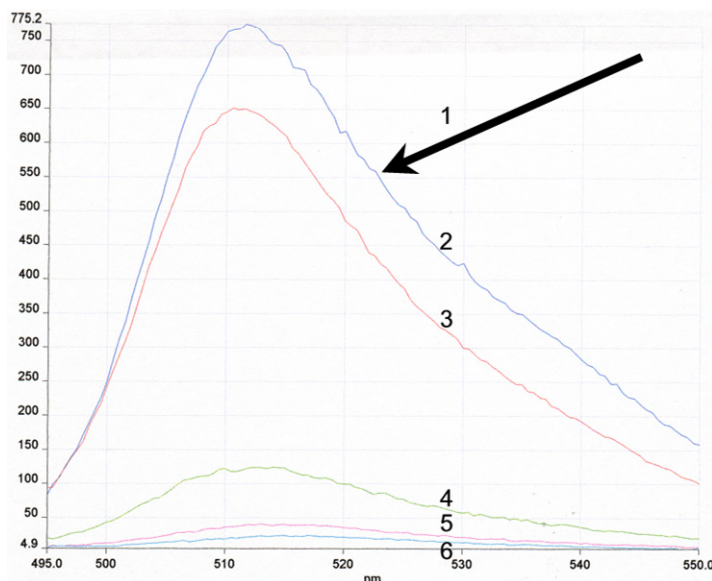
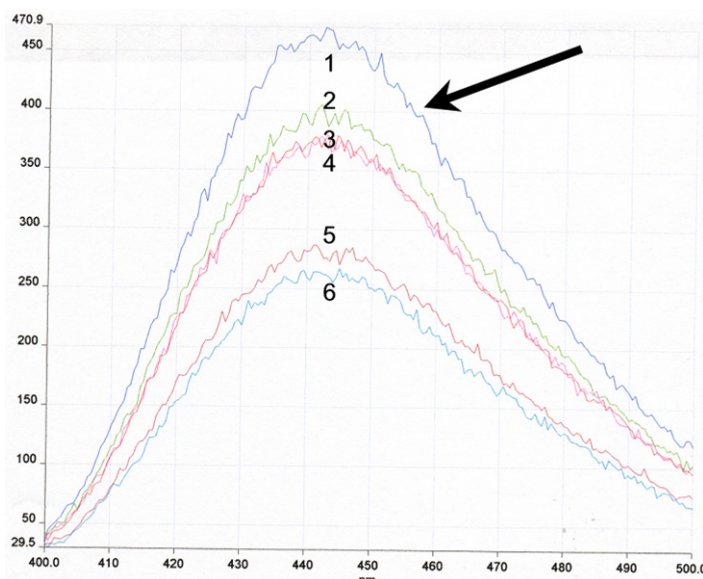


Figure 6-8. Effect of ammonium sulfite on BODIPY labeled nucleotide. 10 μ M BODIPY-GTP γ S is mixed with ammonium sulfite as explained in the text. The black arrow points to the 0 time point; scans are numbered 1-5 for every 10 minute time point. Excitation is at 504 nm and emission for BODIPY is measured from 495-550 nm.

Figure 6-9. Effect of ammonium sulfite on MANT labeled nucleotide. 5 μ M MANT-GTP is mixed with ammonium sulfite as explained in the text. The black arrow points to the 0 time point; scans are numbered 1-6 for every 10 minute time point. Excitation is at 356 nm and emission is measured from 400-500 nm.



It immediately became clear that the ammonium sulfite was having an effect on the stability of the fluorescent signal for both BODIPY and MANT nucleotides. Various control experiments were used to verify this hypothesis, including testing fluorescence in water or buffer alone over the same time frame (data not shown).

These controls showed no loss of fluorescent signal over the hour long experimental time frame. The experiment with the fluorophores bound to protein could not be done, since the complex precipitated in 2M ammonium sulfite as observed in the crystal experiments.

There are two published crystal conditions for untagged $G\alpha_{i1}$ (Coleman, Lee et al. 1994). In an attempt to salvage the crystallographic experiment, the potassium phosphate condition (potassium phosphate, pH 7, 100 mM NaCl) was tested to see if it also produced a deleterious effect on the fluorophore nucleotides.

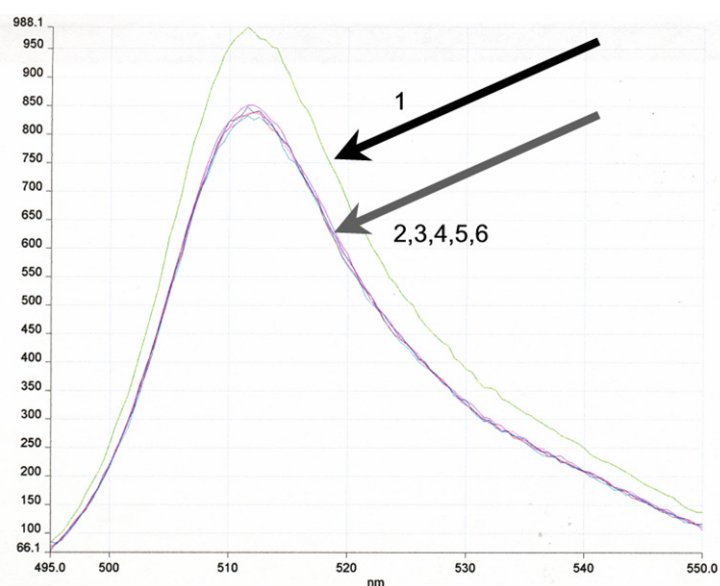


Figure 6-10. Effect of potassium phosphate on BODIPY labeled nucleotide. 5 μ M BODIPY-GTP γ S was mixed with potassium phosphate condition as outlined in the text. The black arrow points to the 0 time point; the grey arrow is pointing to the next five scans, taken every 10 minutes for one hour. Excitation and emission are as in Figure 6-8.

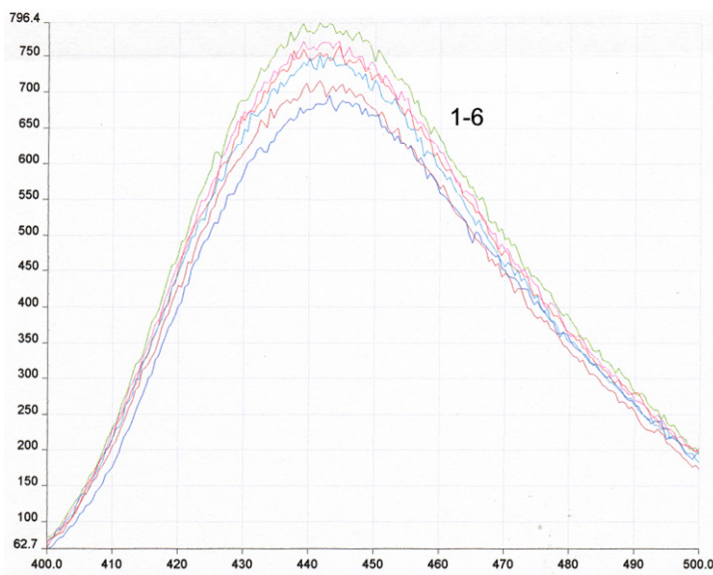


Figure 6-11. Effect of potassium phosphate on MANT-labeled nucleotide. 10 μ M MANT-GTP was mixed with potassium phosphate condition as outlined in the text. The first time point is in green; scans are taken at every 10 minute time point, and the last time point is in purple. Excitation and emission are as in Figure 6-9.

The conclusion drawn from this set of experiments, along with the crystallographic observations, was that the fluorophore• $G_{\alpha_{i1}}$ complex could not be crystallized in the ammonium sulfite condition due to a chemical sensitivity contained in the fluorophore group of the nucleotide. Crystallization trials of the fluorophore• $G_{\alpha_{i1}}$ complex must now either focus on the less reproducible potassium phosphate condition, or new conditions must be found.

Attempted Optimization of Potassium Phosphate Crystallization Condition of the $G_{\alpha_{i1}}$ •Fluorophore Complex

Attempts to crystallize the $G_{\alpha_{i1}}$ protein using the potassium phosphate condition were, on the whole, unsuccessful. This condition is less reproducible than the ammonium sulfite condition; in addition, the crystals do not appear until one month

after being set up, making optimization non-trivial. Crystals were obtained of the GDP-bound form of the modG α_{i1} protein (Figure 6-12); however, GTP γ S bound crystals have been obtained only as microcrystals. In the case of the GDP crystals, more optimization is needed; the crystals are too thin and fragile to be cryoprotected and frozen efficiently and to date do not diffract.

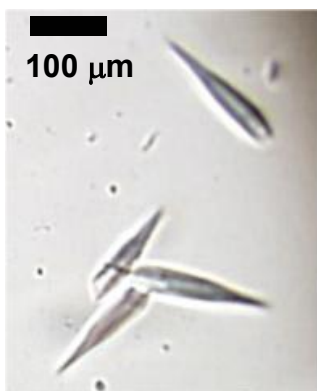


Figure 6-12. Crystals of the GDP·G α_{i1} complex.

Crystal Behavior of the G α_{i1} •BODIPY-GTP γ S Complex

Attempts have been made to crystallize the G α_{i1} •BODIPY-GTP γ S complex using the potassium phosphate conditions. However, only microcrystals have been obtained, and the crystals are too small to test for diffraction.

In conclusion, the G α_{i1} crystals obtained in the potassium phosphate based crystallization condition are less reproducible and more difficult to optimize than the ammonium sulfite condition. However, due to the reactivity of the MANT and BODIPY coupled nucleotides, any further crystallographic studies of these complexes must either be done using potassium phosphate as the precipitant, or a new crystallization for these studies must be found. Perhaps G α_s would be a better

isoform for these co-crystallization studies; $G\alpha_s$ crystallizes in a PEG based precipitant, thereby avoiding the issues encountered in the $G\alpha_{i1}$ crystallization.

BIBLIOGRAPHY

- Coleman, D. E., A. M. Berghuis, et al. (1994). "Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis." Science **265**(5177): 1405-12.
- Coleman, D. E., E. Lee, et al. (1994). "Crystallization and preliminary crystallographic studies of Gi alpha 1 and mutants of Gi alpha 1 in the GTP and GDP-bound states." J Mol Biol **238**(4): 630-4.
- Coleman, D. E. and S. R. Sprang (1998). "Crystal structures of the G protein Gi alpha 1 complexed with GDP and Mg²⁺: a crystallographic titration experiment." Biochemistry **37**(41): 14376-85.
- Coleman, D. E. and S. R. Sprang (1999). "Structure of Gialpha1.GppNHp, autoinhibition in a galpha protein-substrate complex." J Biol Chem **274**(24): 16669-72.
- Mixon, M. B., E. Lee, et al. (1995). "Tertiary and quaternary structural changes in Gi alpha 1 induced by GTP hydrolysis." Science **270**(5238): 954-60.

CHAPTER SEVEN

Conclusions and Future Directions

G β γ Interactions with Peptide/Protein Targets

The structure of SIGK bound to G β γ presented in Chapter Four addresses a number of issues related to G β γ signaling and the nature of the top face of G β . The 2.7 Å structure shows the SIGK peptide binding to the top face of G β using many of the same residues sampled by G β γ binding proteins. This structure allowed for a detailed analysis of the top face of G β , driving a series of mutations to be made that could then be tested using SIGK binding as a probe. This analysis leads to the conclusion that the majority of residues that make up the SIGK binding interface on G β contribute significant energy to the binding of this peptide; that is, individual mutations of G β residues have strong effects on the binding affinity of the whole peptide for G β . Future work along these lines could include testing the SIGK interface mutations for their effect on the other peptides from the phage display assay. This would allow the researcher to understand if there were G β residues crucial for all peptide binding, or if G β residues in the SIGK interface could be used to categorize peptides on the basis of their binding interface. Since the top face of G β is a preferred interaction surface, one would think there would be many different combinations of G β residues that bound to peptides of disparate sequence. In fact, this seems to be the case (Figure 7-1). Perhaps these peptides could then be tested for their ability to affect G β γ target

proteins; one would expect that if the phage display peptides bound to $G\beta$ using different binding determinants, this could translate into differential abilities to block various $G\beta\gamma$ signaling pathways. The ability to selectively block $G\beta\gamma$ signaling through different effectors would be an extremely useful outcome of these experiments.

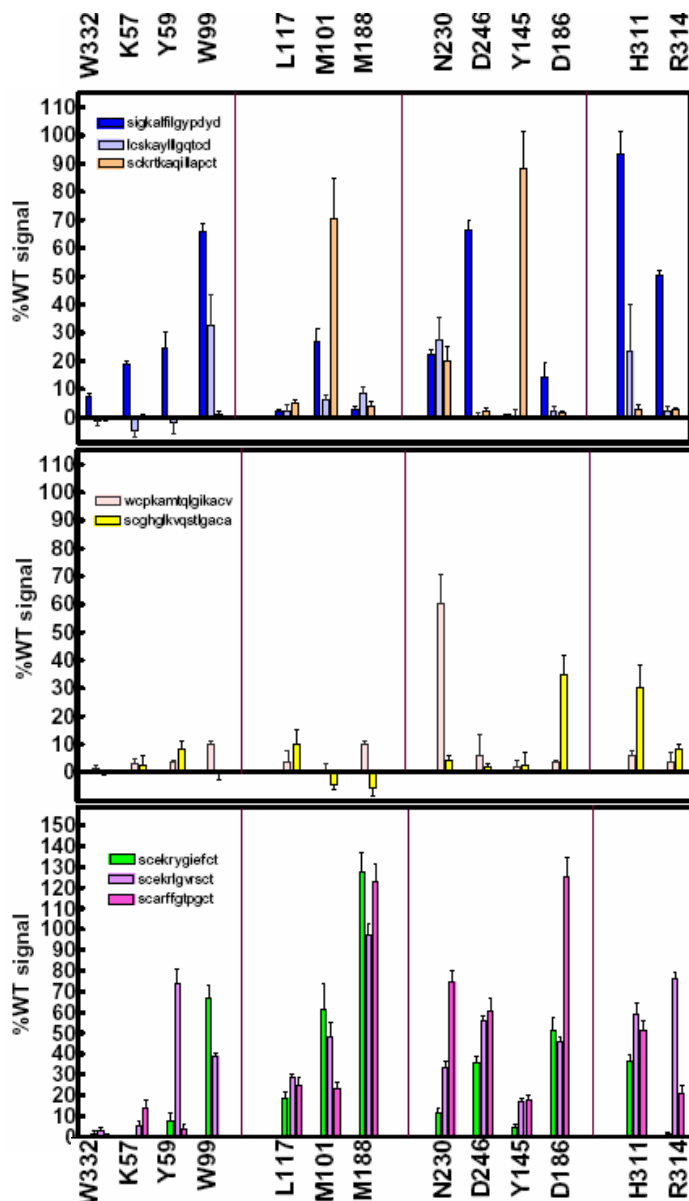


Figure 7-1. Peptides of diverse sequence bind $G\beta$ using distinct residues.

Using the same phage ELISA procedure outlined in Chapter Four, peptides of diverse sequence are tested for their ability to bind to $G\beta$. Surprisingly, even peptides grouped in the same “class” as shown in Figure 2-1 differentially bind to point mutants of $G\beta$, indicating a unique footprint for each peptide shown in regard to $G\beta$ binding. This figure is from unpublished work by T. Bonacci and A. V. Smrcka.

In addition, the SIGK•G $\beta\gamma$ complex led directly to the analysis of the top face of G β as a “hot spot” for protein binding. The conclusion drawn in Chapter Four is that although G β is a functional hot surface containing protein, the properties of G β that allow it to carry out this function are mainly due to the unique composition of the amino acids in the SIGK interface. The combination of nonpolar and aromatic amino acids, along with a large group of acidic charges near the SIGK N terminal interface, allow G β to accommodate proteins with large diversity of sequence and structural elements. The methods currently in the literature used to define hot spot surfaces fail to delineate the top face of G β as a hot surface, although clearly this region of G β is a functional hot spot. The G β mutations designed based on the SIGK•G $\beta\gamma$ structure may now be used to test the interactions of G β with various target binding proteins, including G α , GRK2, phosducin, and adenylyl cyclase. If the mutations to alanine produce a 10-fold or greater loss of affinity, this would provide even stronger evidence for the SIGK binding interface to be described as a new type of “hot surface”. Perhaps then the inclusion of G β properties into hot spot databases would broaden the definition of hot spot proteins and improve the predictive qualities of these databases.

VITAE

Tara Lynne Davis was born in Dallas, Texas, on January 30, 1973. She is the daughter of Barbara Ann and Jerry Davis, and the sister of Steve Davis. She graduated from North Mesquite High School in 1991, and completed a Bachelor of Science in Biochemistry at the University of Texas at Austin in 1995. She was employed as an analytical chemist in the Research and Development department of D&S Plastics for one year, and then entered the Graduate School of Biomedical Sciences at the University of Texas Southwestern Medical Center in Dallas in 1997. She was admitted to candidacy in the Molecular Biophysics program in 1999.

Permanent Address: c/o Steve Davis; 1817 Foxborough Trail
Flower Mound, Texas, 75028